Optimal Pants Decompositions and Shortest Homotopic Cycles on an Orientable Surface*

Éric COLIN DE VERDIÈRE
École normale supérieure, Paris, France

Francis LAZARUS
GIPSA-Lab, Grenoble, France

Abstract

We consider the problem of finding a shortest cycle (freely) homotopic to a given simple cycle on a compact, orientable surface. For this purpose, we use a *pants decomposition* of the surface: a set of disjoint simple cycles that cut the surface into *pairs of pants* (spheres with three holes). We solve this problem in a framework where the cycles are closed walks on the vertex-edge graph of a combinatorial surface that may overlap but do not cross.

We give an algorithm that transforms an input pants decomposition into another homotopic pants decomposition that is *optimal*: each cycle is as short as possible in its homotopy class. As a consequence, finding a shortest cycle homotopic to a given simple cycle amounts to extending the cycle into a pants decomposition and to optimizing it: the resulting pants decomposition contains the desired cycle. We describe two algorithms for extending a cycle to a pants decomposition. All algorithms in this paper are polynomial, assuming uniformity of the weights of the vertex-edge graph of the surface.

1 Introduction

The computation of shortest paths in a geometric domain is one of the fundamental problems in computational geometry. There are many instances of this task, depending on the

*Authors' addresses: É. Colin de Verdière, Laboratoire d'informatique, École normale supérieure, CNRS, Paris, France, e-mail: Eric.Colin.de.Verdiere@ens.fr; F. Lazarus, GIPSA-Lab, CNRS, Grenoble, France, e-mail: Francis.Lazarus@gipsa-lab.inpg.fr. A preliminary version of this paper appeared in *Proceedings of the 11th International Symposium on Graph Drawing, 2003.* Portions of this work were done while the first author was visiting CWI Amsterdam.

underlying topological space (2-manifold, 3-manifold, non-manifold, higher-dimensional space, etc.), the way the lengths are measured in this space (Euclidean distance, L_p distance, etc.), and the possible additional constraints on the solution path. See [Mit00] for a recent and comprehensive survey.

A more specific question is to find a shortest path within a given homotopy class: given a path, find the shortest path that can be obtained from it by continuous deformations, while keeping its endpoints fixed. This is also relevant for loops (closed paths with basepoint) and cycles (without basepoint—this is also called *free* homotopy). In all cases, the underlying space should not be too complicated: indeed, the problem of determining whether a loop is contractible (deformable to a point) is undecidable for 2-simplicial complexes and for 4-manifolds, and open for 3-manifolds [Sti93, p. 242–247]. It is thus natural to consider the class of surfaces, on which the homotopy problem is tractable [DG99]. In this paper, we study the problem of finding a shortest cycle homotopic to a given cycle on a surface. Before describing our results in detail, we present related works.

1.1 Previous Works

A few noteworthy facts are known in the case of smooth surfaces. For hyperbolic surfaces, there is a unique closed geodesic (a cycle that is locally of minimal length) in each homotopy class, which is also a shortest homotopic cycle [Bus92, Theorem 1.6.6]. Any iterative process that locally shortens a cycle will converge to this cycle; this provides an elementary (though non-finite) algorithm [Bus92, Appendix].

For surfaces with a Riemannian metric, [HS94a] analyze a shortening scheme that produces closed geodesics; but this process is not finite and may fail to converge to the shortest geodesic. Also in this case, [FHS82] prove a remarkable result: any simple cycle (a cycle without self-intersection) can be shortened as much as possible in its homotopy class in such a way that the resulting cycle is also simple. Their proof relies on the ability to shorten curves with corners by smoothing and on the fact that geodesics can only cross transversally.

In the case of piecewise-linear surfaces, [HS94b] give an optimal linear-time algorithm (in the real-RAM model) to compute shortest homotopic paths and cycles on a surface made of Euclidean triangles, if all the triangle vertices lie on the surface boundary. Algorithms for the case of the plane minus some obstacles also exist [EKL06, Bes03].

Most of the works that address the computational issues of these problems, however, consider a *combinatorial surface*, described in detail in a former paper [CdVL05]: The surface has a weighted graph embedded on it, so that each face is homeomorphic to an open disk; disjoint simple curves are walks on the graph that may overlap but do not cross (equivalently, they may share edges and vertices, but can always be spread apart so as to become simple and disjoint on the surface); their length is the sum of the weights of the edges traversed, counted with multiplicity. In the present paper, all our results and techniques are done in (a model equivalent to) combinatorial surfaces. This notion has already been used in the context of non-crossing shortest paths in planar graphs [TSN96]; we describe the works that are more related to our topic.

A series of papers consider topological decompositions of combinatorial surfaces. A *cut graph* is a set of curves that cut a surface into a disk; a *system of loops* is a cut graph made of simple loops sharing a common basepoint and otherwise disjoint; a *canonical system of loops* is a system of loops that meet in a particular cyclic order around the vertex. Based on an early paper by [VY90], [LPVV01] give two time-optimal algorithms to compute a canonical system of loops of a combinatorial surface; [EHP04] consider the problem of finding the shortest cut graph; [EW05] describe an algorithm to compute the shortest system of loops of a surface in polynomial time.

A related problem is the computation of a shortest non-contractible or non-separating cycle on a combinatorial surface, which can be seen as an elementary step to build a topological decomposition of a surface. [EHP04] give algorithms to achieve this task using a "circular wave expansion"; very recently, [CM07] and [Kut06] give algorithms with better running-times in some cases. [MT01, Section 4.3, p. 110] describe more general families of cycles for which a wave expansion process will find the shortest cycle.

In a previous paper [CdVL05], we consider the following problem: given a simple loop ℓ on a combinatorial surface, compute the shortest loop that is simple and homotopic to ℓ. We provide an algorithm with polynomial running-time in the case of uniform weights. Key ingredients are the use of a decomposition of the surface by a system of loops and the study of the way some specific curves can cross. In the present paper, the general ideas are similar, but the analysis is more complicated.

1.2 Our Results

In this paper, we consider the problem of finding a shortest cycle homotopic to a given simple cycle on an orientable combinatorial surface (defined above). For this purpose, we will actually shorten a *pants decomposition* [Hat00] of the surface: a set of disjoint simple cycles that cut the surface into *pairs of pants* (spheres with three holes); see Figure 1.

We describe a conceptually simple, greedy process that takes a pants decomposition $s = (s_1, \ldots, s_n)$, where the s_i are the cycles, and outputs a shorter pants decomposition $r = (r_1, \ldots, r_n)$, so that s_i and r_i are homotopic for each i. We prove that each cycle r_i is *optimal*, i.e., as short as possible in its homotopy class. In particular, the resulting decomposition is *optimal* in the sense that it is as short as possible among all pants decompositions made of homotopic cycles, and any optimal pants decomposition is made of optimal cycles. Furthermore, this leads to an algorithm with running-time polynomial in the complexity of the surface and of the input pants decomposition, assuming uniform weights for the edges of the combinatorial surface. [EHP04, Conclusion] raise the problem of shortening a pants decomposition of a combinatorial surface; to our knowledge, we present the first algorithm for this purpose.

Now, if γ is a non-contractible simple cycle on a combinatorial surface, we can compute

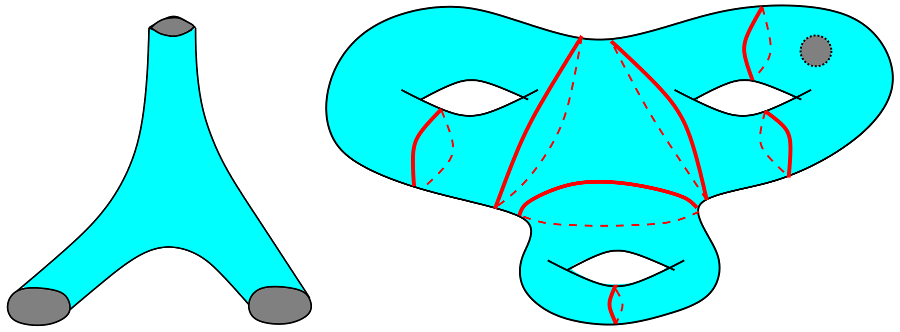

Figure 1: Left: a pair of pants. Right: a pants decomposition of a surface with genus three and one boundary.

a shortest cycle homotopic to γ that has the property of being simple: it suffices to extend γ into a pants decomposition and to optimize it; the resulting pants decomposition contains such a cycle. We provide two methods to extend a simple cycle into a pants decomposition; this implies that this optimization problem has polynomial complexity (in the case of uniform weights), which was previously unknown.

The simplicity of a shortest cycle homotopic to a given simple cycle, in our model as well as in the piecewise-linear case, cannot be proved using the same arguments as in [FHS82], who heavily rely on differential properties of the surface. A proof for those cases follows easily from a result in a subsequent paper by [HS85, Theorem 2.7] stating that a self-intersecting cycle homotopic to a simple cycle must have an embedded 1-gon or 2-gon. While more of a topological nature, their proof uses a non-trivial transversality argument from differential topology.

Assuming the simplicity of a shortest cycle homotopic to a simple cycle, the fact that an optimal pants decomposition is made of optimal cycles can be proved using previous results. Let $(t_1, \ldots, t_n)$ be shortest (simple) cycles homotopic to $(s_1, \ldots, s_n)$. By [HS85, Lemma 3.1], if the cycles $(t_1, \ldots, t_n)$ cross, there must be a 2-gon between them, bounded by two curves that must have equal length and can be uncrossed, decreasing the number of intersections. This result also follows from [dGS97], who prove that any set of cycles can be transformed by *Reidemeister moves* to a set of homotopic cycles with the least possible number of intersections: if we consider the overlay of the optimal pants decomposition $(r_1, \ldots, r_n)$ and some (simple) t_i, the only possible Reidemeister moves consist of "uncrossing" two cycles that cross twice, bounding a 2-gon. While the result by [dGS97] is of a purely topological nature, its proof heavily relies on hyperbolic geometry, in contrast to our paper.

In the present paper, we do not assume the simplicity of a shortest cycle homotopic to a simple cycle; rather, we prove it in the case of combinatorial surfaces. Our proofs are based on the study of crossings between some cycles, though we do not make use of any geometrical or differential property. We do indeed essentially use topological arguments, ultimately relying on the Jordan–Schönflies theorem. In the combinatorial surface model, the metric properties are easier to deal with than for piecewise-linear surfaces, where the structure of shortest paths can be complicated (on such a surface, two shortest paths can overlap for some time without crossing in the topological sense). This allows to point out the topological difficulties of the problem without entangling another involved formalism due to the metric. Nevertheless, our proof techniques could allow for similar results in a more general setting such as cycles drawn on a piecewise-linear surface.

This paper is organized as follows. We review elementary topological notions, describe our optimization process for pants decompositions, and state the optimality theorem in Section 2. The proof of this result is given in the next three sections. We then discuss the computational and complexity issues for optimizing a pants decomposition (Section 6) and a simple cycle (Section 7).

2 Framework and Result

2.1 Topological Background

We begin with some useful definitions from standard topology [Sti93, Hat02, Mas77].

A *surface* $\mathcal{M}$ (possibly with boundary) is a topological Hausdorff space in which each point has a neighborhood homeomorphic to either the plane or the closed half-plane. The points without neighborhood homeomorphic to the plane are the *boundary* of $\mathcal{M}$. In this paper, unless otherwise stated, we only consider *connected*, *compact*, *orientable* surfaces, *possibly with boundary*; such a surface is homeomorphic to a sphere with g handles glued and b boundary disks removed, for some unique integers $g \geq 0$ and $b \geq 0$; g is called the *genus* of $\mathcal{M}$ and b its *number of boundaries*. For example, $g = b = 0$ for the sphere; $g = 0$ and $b = 1$ for the disk; $g = 0$ and $b = 2$ for the cylinder; $g = 0$ and $b = 3$ for the *pair of pants* (see Figure 1, left); $g = 1$ and $b = 0$ for the torus.

Let $\mathcal{M}$ be a connected, compact, orientable surface, possibly with boundary. A *path* on $\mathcal{M}$ is a continuous map $p : [0, 1] \to \mathcal{M}$; its *endpoints* are $p(0)$ and $p(1)$. A *closed path*, or *loop*, is a path whose endpoints coincide. An *arc* is a path intersecting the boundary of $\mathcal{M}$ exactly at its endpoints. A *cycle* is a continuous map $\gamma : S^1 \to \mathcal{M}$, where $S^1 = \mathbb{R}/\mathbb{Z}$ is the standard circle. A path or cycle is *simple* if it is one-to-one; a loop is *simple* if its restriction to $[0, 1)$ is one-to-one. In this paper, a *curve* is a path or a cycle.

As usual, the way the curves are parameterized does not really matter (if p is a path and $\varphi : [0, 1] \to [0, 1]$ is an increasing bijection, we could as well consider $p \circ \varphi$ instead of p). A *subpath* of a path p is the restriction of p to some subsegment of $[0, 1]$, and then reparameterized over $[0, 1]$. The *concatenation* of two paths p and q, with $p(1) = q(0)$, is the path $p \cdot q$ defined by

$$(p \cdot q)(t) = \begin{cases} p(2t) & \text{if } t \leq 1/2; \\ q(2t-1) & \text{if } t \geq 1/2. \end{cases}$$

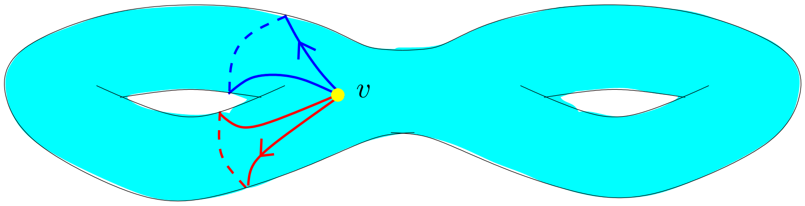

Figure 2: On this double-torus, the two loops with endpoint v are not homotopic; but the corresponding cycles are freely homotopic.

Let $n \in \mathbb{Z}$ and let γ be a cycle. The nth *power* of γ, denoted by γ^n, is the cycle γ iterated n times: $\gamma^n(t) = \gamma(nt \bmod 1)$.

Two paths p and q, both with endpoints a and b, are *homotopic* if there is a continuous family of paths with endpoints a and b that joins p and q. More formally, a *homotopy* between p and q is a continuous map $h : [0,1] \times [0,1] \to \mathcal{M}$ such that $h(0,\cdot) = p$, $h(1,\cdot) = q$, $h(\cdot,0) = a$, and $h(\cdot,1) = b$. A loop is *contractible* if it is homotopic to the constant loop. Two cycles γ and δ are (freely) *homotopic* if there exists a continuous function $h : [0,1] \times S^1 \to \mathcal{M}$ such that $h(0,\cdot) = \gamma$ and $h(1,\cdot) = \delta$.

Let $p : [0,1] \to \mathcal{M}$ be the loop defined by $p(t) = \gamma(t \bmod 1)$. Similarly, let $q : [0,1] \to \mathcal{M}$ be defined by $q(t) = \delta(t \bmod 1)$. The cycles γ and δ are homotopic if and only if there exists a path β joining $p(0)$ to $q(0)$ such that the loop $\beta^{-1} \cdot p \cdot \beta \cdot q^{-1}$ is contractible.

Homotopy of loops (also called homotopy with basepoint) and homotopy of cycles (also called free homotopy) are two different equivalence relations: for instance, two cycles sharing a point v can be freely homotopic, but fail to be homotopic when considered as loops with fixed basepoint v, see Figure 2. A connected space is *simply connected* if every loop in this space is contractible.

A *pants decomposition* of $\mathcal{M}$ is an ordered set of simple, pairwise disjoint cycles that split $\mathcal{M}$ into pairs of pants [Hat00]. Every compact orientable surface, except the sphere, disk, cylinder, and torus, admits a pants decomposition, obtained for example by cutting the surface iteratively along an essential cycle (a simple cycle that does not bound a disk nor a cylinder). Equivalently, a pants decomposition is a maximal ordered set of pairwise disjoint essential cycles such that any two cycles are not freely homotopic. If $\mathcal{M}$ has genus g and b boundaries, a pants decomposition is made of $3g + b - 3$ cycles [Hat00].

Since every cycle is contractible on a sphere or a disk, we exclude these two cases from our study. Furthermore, if $\mathcal{M}$ is a torus or a cylinder, there exists no pants decomposition of $\mathcal{M}$. All the techniques and results used in this paper apply to these surfaces as well (with minor changes) if we allow a pants decomposition to decompose the surface into pairs of pants *and/or cylinders*. For simplicity, however, we will not describe these two particular cases in the sequel.

For technical reasons, it will be easier to work with *doubled pants decompositions*. This is an ordered set of disjoint simple cycles that split $\mathcal{M}$ into cylinders and pairs of pants,

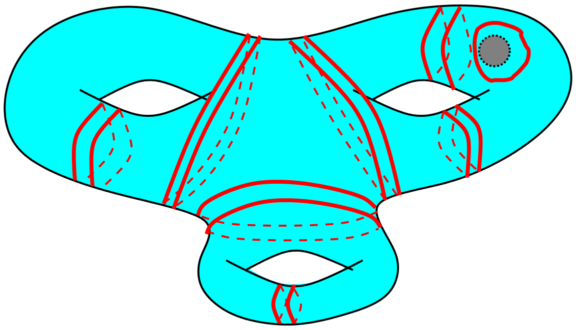

Figure 3: A doubled pants decomposition corresponding to the pants decomposition of Figure 1.

such that each cycle is the boundary of exactly one cylinder and exactly one pair of pants (see Figure 3). Any pants decomposition s of $\mathcal{M}$ can be augmented to a doubled pants decomposition by adding to s a copy of each of its cycles, slightly translated and in the same homotopy class, and a copy of each of the boundary cycles of $\mathcal{M}$, so that the resulting cycles are still simple and pairwise disjoint.

A doubled pants decomposition $s = (s_1, \dots, s_N)$ is made of $N = 6g + 3b - 6$ cycles $s_1, \dots, s_N$. Two cycles of s, or a cycle of s and a boundary of $\mathcal{M}$, that bound a cylinder with no cycle inside it, are called *twins*. Given $j \in [1, N]$, the component of the surface obtained by cutting $\mathcal{M}$ along $s \setminus s_j$ that contains s_j is a pair of pants, denoted by $\mathcal{P}_j$.

2.2 Combinatorial and Cross-Metric Surfaces

Combinatorial surfaces, which we already defined in the introduction, have been used in several papers. In this section, we present an equivalent model (also introduced by [CdVE06]) which simplifies the exposition.

A *cross-metric surface* $\mathcal{M}$ is a surface endowed with an embedded graph $H_{\mathcal{M}}$ such that each open face of $H_{\mathcal{M}}$ is a disk. Each edge of the graph has a positive, possibly infinite, *weight*. Each boundary of $\mathcal{M}$ is the union of some edges of $H_{\mathcal{M}}$, each of them having infinite weight. The sets of curves we consider on $\mathcal{M}$ are sets of curves in the usual sense, except that they must be *regular*, namely:

- the curves are disjoint from the vertices of $H_{\mathcal{M}}$;
- the set of (self-)intersection points of the curves is finite, disjoint from the edges of $H_{\mathcal{M}}$, and each of these points is actually a single (self-)crossing;
- the curves intersect the edges of $H_{\mathcal{M}}$ at finitely many points and actually cross these edges at these points (except for endpoints of arcs).

The *length* $|c|$ of a curve c is the sum of the weights of the edges of $H_{\mathcal{M}}$ crossed by c, counted with multiplicity.

Let $\mathcal{M}$ be a cross-metric surface and C be a set of disjoint simple curves on $\mathcal{M}$. The surface $\mathcal{M}$ cut along C is naturally a cross-metric surface $\mathcal{M}'$: each edge of $H_{\mathcal{M}}$ is split by the curves in C into subedges, forming edges of $H_{\mathcal{M}'}$ with the same weight as the original edge of $H_{\mathcal{M}}$, and each piece of a curve in C cut by $H_{\mathcal{M}}$ corresponds to two boundary edges of $H_{\mathcal{M}'}$ with infinite weight.

In Section 6, we will compute sets of pairwise disjoint simple curves on cross-metric surfaces. For a computational purpose, storing the actual position of the curves on a cross-metric surface $\mathcal{M}$ is unnecessary; all we need is to store the way the considered curves cross the graph $H_{\mathcal{M}}$ (in particular, these informations allow to recover the homotopy classes and lengths of the curves). Said differently, it is sufficient to maintain the combinatorial properties of the *arrangement* of the curves and of the graph $H_{\mathcal{M}}$: the vertices, edges, and faces of the arrangement, and incidence relations between them. The structure of combinatorial map [Tut01, chap. X] is convenient for this purpose. Since the faces of $H_{\mathcal{M}}$ are open disks, and since we only consider non-contractible curves not contained in a face of $H_{\mathcal{M}}$, the faces of the arrangement are also open disks. Curves will be introduced in this arrangement in a *regular* way: each time a curve crosses an edge e of $H_{\mathcal{M}}$, a vertex of degree four is created; e is split into two subedges, each inheriting the weight of e. Furthermore, an endpoint of an arc becomes a vertex of degree three. We can easily compute shortest paths (using Dijkstra's algorithm) or perform breadth-first searches in a cross-metric surface, simply by restating the usual algorithms in the dual graph of $H_{\mathcal{M}}$. In our data structure, we may also simulate cutting $\mathcal{M}$ along some curves by assigning infinite weights to these edges, thus pretending that the edges of these curves become boundary edges and cannot be crossed.

The *complexity* of a cross-metric surface $\mathcal{M}$ is the total number of vertices, edges, and faces of $H_{\mathcal{M}}$. The *complexity* of a curve on a cross-metric surface $\mathcal{M}$ is the number of crossings of the curve with $H_{\mathcal{M}}$.

Most of the results in this paper will be stated in terms of cross-metric surfaces. But we now show that a combinatorial surface $\mathcal{M}$ with weighted vertex-edge graph $G_{\mathcal{M}}$ (as defined in an earlier paper [CdVL05]) can be viewed as a cross-metric surface, so our results apply for combinatorial surfaces as well. Define the embedded graph $G^*_{\mathcal{M}}$, which is a slight variant of the dual graph, as follows (see Figure 4):

- in the interior of each face f of $G_{\mathcal{M}}$, there is a vertex f^* of $G^*_{\mathcal{M}}$;
- in the relative interior of each boundary edge e of $G_{\mathcal{M}}$, there is a vertex $\overline{e}^*$ of $G^*_{\mathcal{M}}$;
- for each non-boundary edge e of $G_{\mathcal{M}}$, with incident faces f_1 and f_2, there is an edge e^* in $G^*_{\mathcal{M}}$ between f_1^* and f_2^* crossing e and no other edge of $G_{\mathcal{M}}$;
- for each boundary edge e, with incident face f, there is an edge e^* in $G^*_{\mathcal{M}}$ between f^* and $\overline{e}^*$ crossing no edge of $G_{\mathcal{M}}$;

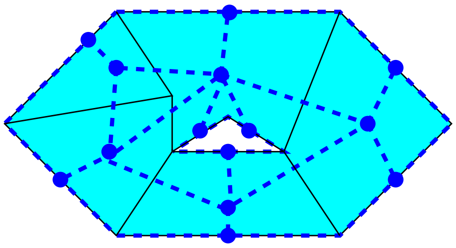

Figure 4: In solid lines: the vertex edge-graph $G_{\mathcal{M}}$ of a combinatorial surface $\mathcal{M}$ with boundary. In dashed lines: the graph $G^*_{\mathcal{M}}$.

- each pair of consecutive vertices of $G^*_{\mathcal{M}}$ on a boundary of $\mathcal{M}$ is connected by an edge of $G^*_{\mathcal{M}}$ on the boundary of $\mathcal{M}$.

Define the weight of e^* to be the length of e, and the weight of the boundary edges to be infinite. The surface $\mathcal{M}$, together with the embedded graph $H_{\mathcal{M}} = G^*_{\mathcal{M}}$, is a cross-metric surface. Any walk w in $G_{\mathcal{M}}$, with sequence of edges $e_1, \ldots, e_k$, corresponds to a path p drawn on $\mathcal{M}$ crossing edges $e^*_1, \ldots, e^*_k$ of $G^*_{\mathcal{M}}$; the length of w (viewed as walk in the graph $G_{\mathcal{M}}$) is the same as the length of the path p (in the cross-metric surface $\mathcal{M}$). Conversely, any path on $\mathcal{M}$ avoiding the vertices of $H_{\mathcal{M}}$ is homotopic to a walk in $G_{\mathcal{M}}$ with the same length. A curve on a combinatorial surface is simple if and only if it can be represented as a simple curve on the corresponding cross-metric surface. As a result, it is sufficient to work with cross-metric surfaces.

2.3 Our Result

Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Let g be its genus and b its number of boundaries; let $N = 6g + 3b - 6$. For $j \in [1, N]$, we define a map f_j that transforms a doubled pants decomposition $s = (s_1, \ldots, s_N)$ of $\mathcal{M}$ into another one, $r = (r_1, \ldots, r_N)$, as follows. If $k \neq j$, then $r_j = s_j$. Furthermore, consider the pair of pants $\mathcal{P}_j$ of $\mathcal{M}$ cut along $s \setminus s_j$ that contains s_j; then r_j is defined to be a shortest cycle, among all simple cycles homotopic to s_j in $\mathcal{P}_j$. The maps f_j are called *elementary steps*. A *main phase* is the map $f = f_N \circ f_{N-1} \circ \ldots \circ f_2 \circ f_1$.[1]

Clearly, these maps transform a doubled pants decomposition into another one. Furthermore, if $r = f_j(s)$, then s_k and r_k are homotopic for each k. Here is our main theorem, proved in the next three sections:

Theorem 2.1 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Let s^0 be a doubled pants decomposition of $\mathcal{M}$, and let $(s^n)_{n\in\mathbb{N}}$ be the sequence defined by $s^{n+1} = f(s^n)$. For each n, write $s^n = (s^n_1, \ldots, s^n_N)$.*

[1]The maps $f_1, \ldots, f_N$ could actually be composed in any order.

For some $m \in \mathbb{N}$, s^m and s^{m+1} have the same length. For such a choice of m, s^m is a doubled pants decomposition such that, for each $i \in [1, N]$, s_i^m is a shortest cycle among all (not necessarily simple) cycles (freely) homotopic to s_i^0.

In particular, s^m is an optimal doubled pants decomposition of $\mathcal{M}$ and contains an optimal pants decomposition.

Moreover, any non-contractible simple cycle is either essential or homotopic to a boundary of $\mathcal{M}$. Such a cycle can be extended to a doubled pants decomposition of $\mathcal{M}$. By Theorem 2.1, we obtain immediately:

Corollary 2.2 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus; let γ be a non-contractible simple cycle in $\mathcal{M}$. There exists a* simple *cycle homotopic to γ and as short as possible among all (not necessarily simple) cycles (freely) homotopic to γ.*

Similar results also hold for the cylinder and the torus: this can be proved using the same techniques, allowing a pants decomposition to split the surface into pairs of pants and/or cylinders. We omit the details.

3 Crossing Words

In this section, we introduce the main ingredient of this paper: the crossing word between a set of disjoint, simple paths or cycles, and a given path or cycle.

3.1 Universal Cover and Lifts

We refer the reader to classical textbooks in algebraic topology [Sti93, Hat02, Mas77] for the details.

Let $\mathcal{M}$ and $\mathcal{N}$ be two possibly non-compact surfaces. A continuous function $\pi : \mathcal{N} \to \mathcal{M}$ is called a *covering map* or *projection* if each point $x \in \mathcal{M}$ lies in an open connected neighborhood U such that $\pi^{-1}(U)$ is a countable union of disjoint open sets $U_1 \cup U_2 \cup \ldots$ and $\pi|U_i : U_i \to U$ is a homeomorphism for each i. If such a map exists, then $\mathcal{N}$ is called a *covering space* of $\mathcal{M}$, and we say that $\mathcal{N}$ *covers* $\mathcal{M}$. Up to a covering isomorphism (a homeomorphism that respects projections), every connected surface $\mathcal{M}$ has a unique simply connected covering space, called the *universal cover* of $\mathcal{M}$ and denoted by $\widetilde{\mathcal{M}}$.

Fix a covering map $\pi : \widetilde{\mathcal{M}} \to \mathcal{M}$. A *translation* (or *automorphism*) of the covering map is a continuous function $\tau : \widetilde{\mathcal{M}} \to \widetilde{\mathcal{M}}$ such that $\pi \circ \tau = \pi$. A *lift* of a path $p : [0, 1] \to \mathcal{M}$ is a path $\widetilde{p} : [0, 1] \to \widetilde{\mathcal{M}}$ such that $\pi \circ \widetilde{p} = p$. Similarly, a lift of a cycle $\gamma : S^1 \to \mathcal{M}$ is a continuous function $\widetilde{\gamma} : \mathbb{R} \to \widetilde{\mathcal{M}}$ such that $\pi(\widetilde{\gamma}(t)) = \gamma(t \bmod 1)$ for all t. We use the following properties of lifts and universal covers.

- The *lift property*: let p be a path in $\mathcal{M}$ with source point x; let $\widetilde{x} \in \widetilde{\mathcal{M}}$ be such that $\pi(\widetilde{x}) = x$. Then there is a unique path $\widetilde{p}$ in $\widetilde{\mathcal{M}}$, starting at $\widetilde{x}$, such that $\pi \circ \widetilde{p} = p$;

- the *homotopy property*: two paths p and q with the same endpoints are homotopic in $\mathcal{M}$ if and only if they have two lifts $\widetilde{p}$ and $\widetilde{q}$ with the same endpoints in $\widetilde{\mathcal{M}}$;
- the *intersection property*: a path p in $\mathcal{M}$ self-intersects if and only if either a lift of p self-intersects, or there exist two lifts of p that intersect each other.

Let c be a simple curve that is either an arc or a non-contractible cycle on $\mathcal{M}$. A *geometric lift* of c is a curve that is a connected component of $\pi^{-1}(c) \subset \widetilde{\mathcal{M}}$. If c is an arc, the geometric lifts of c are exactly the lifts of c. If c is a non-contractible cycle, each geometric lift of c corresponds to infinitely many lifts of c: two lifts $\widetilde{c}$ and $\widetilde{c}'$ of c give rise to the same geometric lift if and only if the image sets of $\widetilde{c}$ and $\widetilde{c}'$ are the same; equivalently, there exists $k \in \mathbb{Z}$ such that $\widetilde{c}(\cdot) = \widetilde{c}'(k + \cdot)$.

Lemma 3.1 *Let $\mathcal{M}$ be a surface and c be either a non-contractible simple cycle in the interior of $\mathcal{M}$ or a simple arc in $\mathcal{M}$. Then each geometric lift of c separates $\widetilde{\mathcal{M}}$ into two connected components.*

This lemma can be proved, in the case where c is a cycle, using hyperbolic geometry [Bus92, p. 417]. In the case where c is an arc, it follows from the fact that the universal covering space of $\mathcal{M}$ is homeomorphic to the unit disk with some boundary points removed [CdVL05, Lemma 3].

Until the end of Section 3, we use the following notations. Let $\mathcal{M}$ be a cross-metric surface. Let C be a set of pairwise disjoint simple curves, each of which is either an arc in $\mathcal{M}$ or a non-contractible cycle in the interior of $\mathcal{M}$. For $c \in C$, the geometric lifts of c are denoted by c^α, where $\alpha \in \mathbb{N}$.[2] The set of all the geometric lifts of the curves in C is denoted by $\widetilde{C}$. By the previous lemma, each geometric lift in $\widetilde{C}$ separates $\widetilde{\mathcal{M}}$ into two connected components.

3.2 Crossing Words for Paths

We consider *words* on the *alphabet* made of letters of the form c^α or $\overline{c}^\alpha$, where $c^\alpha \in \widetilde{C}$. A word y is a *subword* of a word w if w can be written as the concatenation of x, y, and z, where x and z are (possibly empty) words. If a word w contains a subword $c^\alpha\overline{c}^\alpha$ or $\overline{c}^\alpha c^\alpha$, let x be the word resulting from removing this subword from w; we say that x is deduced from w by an *elementary c-reduction.* An *elementary reduction* is an elementary c-reduction for some c. A word w is *(c-)irreducible* if no elementary (c-)reduction can be applied to w. A word w *(c-)reduces* to x if x can be obtained from w by successive elementary (c-)reductions. A word is *parenthesized* if it reduces to the empty word ε.

[2] A natural index set can be described as follows. Pick a basepoint x for both $\mathcal{M}$ and c. Let Γ be the infinite cyclic group generated by the homotopy class of c considered as a loop with basepoint x. The geometric lifts of c are indexed by the set of left cosets of Γ in $\pi_1(\mathcal{M}, x)$. For our purpose, however, it is sufficient (and simpler) to choose an indexation with the set of integers.

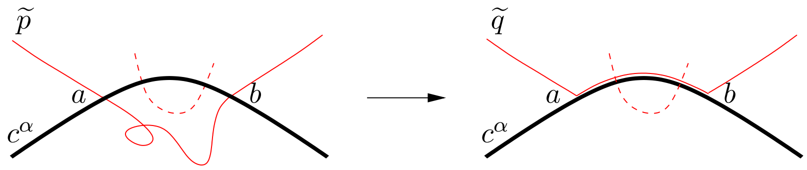

Figure 5: The fundamental operation of uncrossing the parts of two curves c^α and $\widetilde{p}$ between two points a and b corresponding to an elementary reduction on $\widetilde{C}/\widetilde{p}$. The part of $\widetilde{p}$ between a and b is not necessarily simple; also, the part of c^α between a and b can cross other pieces of $\widetilde{p}$.

Let $\widetilde{p}$ be a path in $\widetilde{\mathcal{M}}$. Walk along $\widetilde{p}$ and, at each crossing encountered with a geometric lift $c^\alpha \in \widetilde{C}$, write down the symbol c^α or $\overline{c}^\alpha$, according to the orientation of the crossing (with respect to fixed orientations of $\widetilde{\mathcal{M}}$ and the c^α). Recall that the elements of $\widetilde{C}$ are simple and pairwise disjoint, since the curves in C are simple and pairwise disjoint. The word we obtain is called the *crossing word* of $\widetilde{p}$ with $\widetilde{C}$, and denoted by $\widetilde{C}/\widetilde{p}$.

In all this paper, the following situation will often occur: $\widetilde{p}$ is a lift of a path p, and an elementary reduction is possible on $\widetilde{C}/\widetilde{p}$. This reduction corresponds to two intersection points, a and b, of some geometric lift $c^\alpha \in \widetilde{C}$ with $\widetilde{p}$. The subpaths associated with this possible elementary reduction are the parts of c^α and $\widetilde{p}$ that are between a and b. See Figure 5. We will often replace the part of $\widetilde{p}$ between a and b by a path going along the part of c^α between a and b; we obtain a new path $\widetilde{q}$, lift of some path q on $\mathcal{M}$, which has exactly two crossings fewer with c^α than $\widetilde{p}$ has. Obviously, $\widetilde{C}/\widetilde{q}$ is deduced from $\widetilde{C}/\widetilde{p}$ by proceeding to the elementary reduction; and the new path q is homotopic to p.

The following result is well-known and is due to the fact that the set of words modulo elementary reductions is a free group on $\widetilde{C}$.

Lemma 3.2 *Any word reduces to exactly one irreducible word.*

The following lemma, which relies on the fact that each geometric lift in $\widetilde{C}$ is separating, has been proved in a former paper [CdVL05, Lemma 5]; we include the proof here for completeness.

Lemma 3.3 *Let ℓ be a contractible loop on $\mathcal{M}$, and let $\widetilde{\ell}$ be a lift of ℓ in $\widetilde{\mathcal{M}}$. Then $\widetilde{C}/\widetilde{\ell}$ is parenthesized.*

Proof. We prove the result by induction on the number of crossings between $\widetilde{\ell}$ and the geometric lifts in $\widetilde{C}$. The lemma is trivial if $\widetilde{\ell}$ crosses no element of $\widetilde{C}$. Assume on the contrary there is at least one crossing between $\widetilde{\ell}$ and a geometric lift c^α in $\widetilde{C}$.

Since ℓ is contractible, $\widetilde{\ell}$ is a loop in $\widetilde{\mathcal{M}}$. By Lemma 3.1, $\widetilde{\ell}$ must cross c^α once more with the opposite orientation. We now view $\widetilde{\ell}$ as a cycle $\widetilde{\gamma}$ (i.e., we forget the basepoint of $\widetilde{\ell}$). The two crossings split $\widetilde{\gamma}$ into two paths $\widetilde{p}$ and $\widetilde{q}$ (Figure 6). It is possible to extend $\widetilde{p}$ to a loop $\widetilde{p}'$ so that $\widetilde{C}/\widetilde{p}' = \widetilde{C}/\widetilde{p}$; similarly, we can extend $\widetilde{q}$ to a loop $\widetilde{q}'$ so that

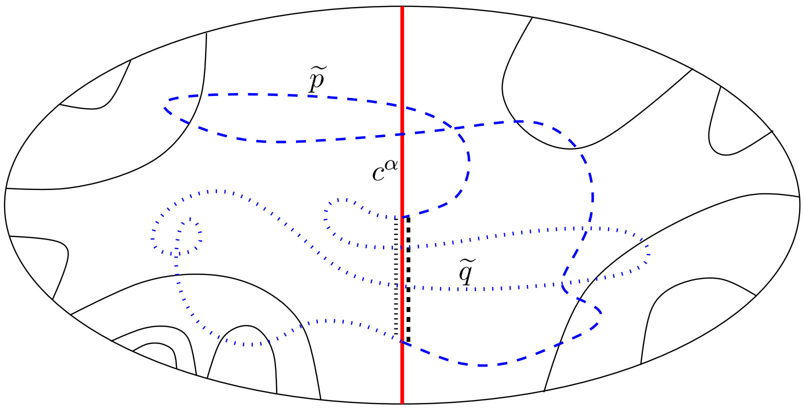

Figure 6: The induction step of Lemma 3.3.

$\widetilde{C}/\widetilde{q}' = \widetilde{C}/\widetilde{q}$. By the induction hypothesis, $\widetilde{C}/\widetilde{p}$ and $\widetilde{C}/\widetilde{q}$ are parenthesized. But $\widetilde{C}/\widetilde{\ell}$ results from the concatenation of these two words, with two additional symbols corresponding to the crossings with c^α, followed by a cyclic permutation: it is thus parenthesized. $\square$

We obtain immediately:

Corollary 3.4 *Let p and q be homotopic paths in $\mathcal{M}$, and let $\widetilde{p}$ and $\widetilde{q}$ be lifts of p and q in the universal cover of $\mathcal{M}$, sharing the same endpoints. Then $\widetilde{C}/\widetilde{p}$ and $\widetilde{C}/\widetilde{q}$ reduce to the same irreducible word.*

3.3 Crossing Word Sets for Cycles

The goal of this section is to define the analogue of the crossing word for the crossings between a lift $\widetilde{\gamma}$ of a non-contractible cycle γ and $\widetilde{C}$.

Any path $\widetilde{\ell} : [0,1] \to \widetilde{\mathcal{M}}$ defined by $\widetilde{\ell}(\cdot) = \widetilde{\gamma}(a + \cdot)$ for some $a \in \mathbb{R}$ such that $\widetilde{\gamma}(a)$ does not belong to a curve in $\widetilde{C}$ is called a *lifted period* of $\widetilde{\gamma}$. The *crossing word set* of $\widetilde{\gamma}$ with $\widetilde{C}$, denoted by $[\widetilde{C}/\widetilde{\gamma}]$, is the set of crossing words $\widetilde{C}/\widetilde{\ell}$, over all lifted periods $\widetilde{\ell}$ of $\widetilde{\gamma}$. Our first task will be to show that the crossing word set $[\widetilde{C}/\widetilde{\gamma}]$ is entirely determined once we know one of its elements.[3]

We note that $\widetilde{\gamma}$ induces a translation $\tau_{\widetilde{\gamma}}$ in $\widetilde{\mathcal{M}}$, as follows. Let $v^0 \in \widetilde{\mathcal{M}}$. Let $\widetilde{\ell}$ be a lifted period of $\widetilde{\gamma}$; consider a path β^0 joining $\widetilde{\ell}(0)$ to v^0; call β^1 the lift of $\pi(\beta^0)$ starting at $\widetilde{\ell}(1)$. The target v^1 of β^1 satisfies $\pi(v^0) = \pi(v^1)$; intuitively, $\widetilde{\gamma}$ translates v^0 to v^1. It is readily seen that v^1 does not depend on the choice of β^0 and $\widetilde{\ell}$. We therefore define $\tau_{\widetilde{\gamma}}(v^0) := v^1$. In particular, $\tau_{\widetilde{\gamma}}$ sends a geometric lift of a curve $c \in C$ to a geometric lift of c.

[3] An alternative definition of the crossing word could use the lift of γ in the covering of $\mathcal{M}$ with cyclic fundamental group generated by this lift. This would avoid the "multiform" of the crossing word but may not simplify the presentation.

For any lift $\widetilde{\gamma}$ and any crossing word w, we define a new crossing word $\varphi_{\widetilde{\gamma}}(w)$ as follows:

$$\varphi_{\widetilde{\gamma}}(w) = \begin{cases} \varepsilon & \text{if } w = \varepsilon, \\ x \cdot \overline{\tau_{\widetilde{\gamma}}(c^\alpha)} & \text{if } w = c^\alpha \cdot x \text{ for some word } x, \\ x \cdot \tau_{\widetilde{\gamma}}(c^\alpha) & \text{if } w = \overline{c^\alpha} \cdot x \text{ for some word } x. \end{cases}$$

(Here, $x \cdot y$ denotes the concatenation of words x and y.) Intuitively, if w is the crossing word $\widetilde{C}/\widetilde{\ell}$ for some lifted period $\widetilde{\ell}$ of $\widetilde{\gamma}$, then $\varphi_{\widetilde{\gamma}}(w)$ is the crossing word of a lifted period of $\widetilde{\gamma}$ shifted from $\widetilde{\ell}$ to overpass exactly one crossing. If w is a word, we define $\langle w \rangle_{\widetilde{\gamma}}$ to be the set $\{\varphi^n_{\widetilde{\gamma}}(w), n \in \mathbb{Z}\}$.

Proposition 3.5 *For any word w in $[\widetilde{C}/\widetilde{\gamma}]$, we have: $[\widetilde{C}/\widetilde{\gamma}] = \langle w \rangle_{\widetilde{\gamma}}$.*

Proof. First, let a, b be two real numbers such that $a < b < a+1$ and exactly one crossing occurs between all geometric lifts of $\widetilde{C}$ and $\widetilde{\gamma}|_{[a,b]}$. Let $w = \widetilde{C}/\widetilde{\gamma}|_{[a,a+1]}$ and $x = \widetilde{C}/\widetilde{\gamma}|_{[b,b+1]}$. We have $x = \varphi_{\widetilde{\gamma}}(w)$; indeed,

$$\begin{aligned} \tau_{\widetilde{\gamma}}(c^\alpha \cap \widetilde{\gamma}|_{[a,b]}) &= \tau_{\widetilde{\gamma}}(c^\alpha) \cap \tau_{\widetilde{\gamma}}(\widetilde{\gamma}|_{[a,b]}) \\ &= \tau_{\widetilde{\gamma}}(c^\alpha) \cap \widetilde{\gamma}|_{[a+1,b+1]}. \end{aligned}$$

From this fact, it is easy to conclude. $\square$

The sets of words of the form $\langle w \rangle_{\widetilde{\gamma}}$ are called the *$\widetilde{\gamma}$-word sets.* Note that $\varphi_{\widetilde{\gamma}}$ does not affect the length of a word, so the *length* of a $\widetilde{\gamma}$-word set is well-defined. Let W and X be $\widetilde{\gamma}$-word sets. If, for some words $w \in W$ and $x \in X$ and for some curve c in C, w elementarily c-reduces to x, we say that W *elementarily c-reduces* to X. When an elementary c-reduction is possible on $[\widetilde{C}/\widetilde{\gamma}]$, exactly the same phenomenon occurs as in Figure 5 (with $\widetilde{\gamma}$ instead of $\widetilde{p}$), and we may also proceed to the reduction by modifying γ, removing the two crossings.

Lemma 3.6 *Any $\widetilde{\gamma}$-word set W c-reduces to exactly one c-irreducible $\widetilde{\gamma}$-word set. Any $\widetilde{\gamma}$-word set W reduces to exactly one irreducible $\widetilde{\gamma}$-word set.*

Proof. The proof is based on a confluence property on reductions of $\widetilde{\gamma}$-word sets. We prove the result for reductions, the same argument holds for c-reductions. Let w be a word; a *simplification* on w consists of either an elementary reduction on w (removal of $c^\alpha \overline{c}^\alpha$ or $\overline{c}^\alpha c^\alpha$), or in the removal of the first and last symbols of w, if the first is of the form c^α or $\overline{c}^\alpha$ and the last of the form $\overline{\tau_{\widetilde{\gamma}}(c^\alpha)}$ or $\tau_{\widetilde{\gamma}}(c^\alpha)$, respectively. It is easily proved that W elementarily reduces to X if and only if, for *any* $w \in W$, there exists $x \in X$ such that w simplifies to x.

We say that two words w and x are *equivalent* if $\langle w \rangle_{\widetilde{\gamma}} = \langle x \rangle_{\widetilde{\gamma}}$. Let w be a word; suppose that w_1 and w_2 are derived from w by a simplification. It can be shown by an easy case analysis that there exist equivalent words x_1 and x_2 such that, for $i = 1, 2$, x_i is derived from w_i by zero or one simplification.

Let W be any $\widetilde{\gamma}$-word set. Assume that W elementarily reduces to W_1 and W_2. Let $w \in W$; by the first paragraph, there exist w_1 and w_2 in W_1 and W_2 such that w simplifies to w_1 and w_2. It follows, by the previous paragraph, that there exists X deduced from W_1 and W_2 by zero or one elementary reduction.

We can now prove the result by induction on the length of $\widetilde{\gamma}$-word sets; the lemma is trivial if the length is 0 or 1. Assume that the lemma is true for all $\widetilde{\gamma}$-word sets of length at most n; Let W be a $\widetilde{\gamma}$-word set of length $n+1$. If W is reducible, consider any two $\widetilde{\gamma}$-word sets, W_1 and W_2, derived from W by an elementary reduction. By the preceding paragraph, W_1 and W_2 reduce (with zero or one elementary reduction) to some X. From our induction hypothesis, each of W_1, W_2, and X reduces to only one irreducible $\widetilde{\gamma}$-word set, which must hence be the same. This concludes the proof. $\square$

We define $g_c^{\widetilde{\gamma}}(W)$ to be the unique c-irreducible $\widetilde{\gamma}$-word set to which W c-reduces. Similarly, $g^{\widetilde{\gamma}}(W)$ is the unique irreducible $\widetilde{\gamma}$-word set to which W reduces.

Proposition 3.7 *Let γ be a cycle homotopic in $\mathcal{M}$ to some cycle δ disjoint from C. Let $\widetilde{\gamma}$ be a lift of γ. Then $g^{\widetilde{\gamma}}([\widetilde{C}/\widetilde{\gamma}]) = \{\varepsilon\}$.*

Proof. Let $p : [0,1] \to \mathcal{M}$ be the loop defined by $p(t) = \gamma(t \bmod 1)$; similarly, let $q : [0,1] \to \mathcal{M}$ be defined by $q(t) = \delta(t \bmod 1)$. There exists a path β joining $p(0)$ to $q(0)$ such that the path $r := \beta^{-1} \cdot p \cdot \beta \cdot q^{-1}$ is contractible in $\mathcal{M}$. Let $\widetilde{r}$ be a lift of r, concatenation of the inverse of β^0, $\widetilde{p}$, β^1, and the inverse of $\widetilde{q}$ (respectively lifts of β, p, β, and q). We choose $\widetilde{r}$ so that $\widetilde{p}$ is a lifted period of $\widetilde{\gamma}$.

Since q is disjoint from C, the word $w := \widetilde{C}/\widetilde{r}$ is the concatenation of $\widetilde{C}/(\beta^0)^{-1}$, $\widetilde{C}/\widetilde{p}$, and $\widetilde{C}/\beta^1$. Furthermore, $\tau_{\widetilde{\gamma}}(\beta^0)$ is equal to β^1; hence, if the kth symbol of $\widetilde{C}/\beta^0$ is equal to c^α (resp. $\overline{c}^\alpha$), then the kth symbol of $\widetilde{C}/\beta^1$ is equal to $\tau_{\widetilde{\gamma}}(c^\alpha)$ (resp. $\overline{\tau_{\widetilde{\gamma}}(c^\alpha)}$). It follows that $\langle w \rangle_{\widetilde{\gamma}}$ reduces to $\langle \widetilde{C}/\widetilde{p} \rangle_{\widetilde{\gamma}} = [\widetilde{C}/\widetilde{\gamma}]$. Now, by Lemma 3.3, w is parenthesized, so $\langle w \rangle_{\widetilde{\gamma}}$ also reduces to $\{\varepsilon\}$. Lemma 3.6 concludes. $\square$

4 Optimal Curves on Pairs of Pants

In this section, we use the crossing word techniques to prove some basic facts regarding curves on cylinders or pairs of pants.

4.1 Optimal Cycles on Pairs of Pants

Proposition 4.1 *Let $\mathcal{K}$ be a cross-metric surface that is a cylinder or a pair of pants. Let ν be a boundary of $\mathcal{K}$. There exists a cycle that is as short as possible among all cycles homotopic to ν and that has the property of being simple.*

The proof relies on the two following lemmas.

Lemma 4.2 *Let $\mathcal{P}$ be a cross-metric pair of pants. Let ν be a boundary of $\mathcal{P}$ and let γ be a cycle homotopic to ν. Let q be a shortest path between two points, one on each of the two boundaries of $\mathcal{P}$ different from ν. There exists a cycle δ that is homotopic to γ, no longer than γ, and that does not cross q.*

PROOF. Let $\widetilde{Q}$ be the set of geometric lifts of q, and let $\widetilde{\gamma}$ be a lift of γ. $[\widetilde{Q}/\widetilde{\gamma}]$ reduces to $\{\varepsilon\}$ by Proposition 3.7 applied on $\mathcal{P}$. If this crossing word is not empty, let γ_1 and q_1 be the homotopic subpaths of γ and q corresponding to an elementary reduction. Since q is a shortest path, $|q_1| \leq |\gamma_1|$, and we can, like in Figure 5, proceed to the elementary reduction by changing γ to another cycle homotopic to and no longer than γ, and having exactly two crossings fewer than γ with q. The proof is finished by induction on the number of crossings between γ and q. $\square$

Lemma 4.3 *Let $\mathcal{C}$ be a cross-metric cylinder. Let γ be a cycle homotopic to the boundaries of $\mathcal{C}$. Let r be a shortest path between two points, one on each of the two boundaries of $\mathcal{C}$. There exists a cycle δ that is homotopic to γ, no longer than γ, and that crosses r exactly once.*

PROOF. Let $\widetilde{R}$ be the set of geometric lifts of r, and let $\widetilde{\gamma}$ be a lift of γ. Since γ "winds around" $\mathcal{C}$ exactly once, the $\widetilde{\gamma}$-word set $[\widetilde{R}/\widetilde{\gamma}]$ reduces to a $\widetilde{\gamma}$-word set of length one, hence has odd length. If γ crosses r at least twice, an elementary reduction is possible on $[\widetilde{R}/\widetilde{\gamma}]$. Let γ_1 and r_1 be the homotopic subpaths of γ and r corresponding to an elementary reduction. Like in Figure 5, we can modify γ to get a no longer, homotopic cycle that has exactly two fewer crossings with r. We finish by induction on the length of $[\widetilde{R}/\widetilde{\gamma}]$. $\square$

PROOF OF PROPOSITION 4.1. We consider any cycle γ homotopic to ν and prove that there exists a homotopic cycle that is simple and no longer. Assume first $\mathcal{K}$ is a pair of pants. By Lemma 4.2, there exists a cycle δ, homotopic to γ, no longer than γ, and that does not cross a shortest path q between the two boundaries of $\mathcal{K}$ that are different from ν. Let $\mathcal{C}$ be the cylinder obtained by cutting $\mathcal{K}$ along q.

The cycle δ is homotopic in $\mathcal{C}$ to the boundaries of $\mathcal{C}$. Indeed, δ is homotopic in $\mathcal{K}$ to ν, and δ is homotopic to some power ν^k of ν in $\mathcal{C}$, hence also in $\mathcal{K}$. It follows that ν and ν^k are homotopic in $\mathcal{K}$, which easily implies that $k = 1$, whence the result. To conclude this case, it now suffices to prove that there exists a cycle that is no longer than δ, homotopic to δ in $\mathcal{C}$, and that is simple; in other words, it suffices to prove our result when $\mathcal{K}$ is a cylinder.

So, let now $\mathcal{K}$ be a cylinder. By Lemma 4.3, we may assume that γ crosses a shortest path r between the boundaries of $\mathcal{K}$ exactly once, say at some point a. Cut $\mathcal{K}$ along r; γ becomes a path whose endpoints are the two copies a' and a'' of a. Hence, a shortest path between a' and a'' leads, after regluing along r, to a cycle in $\mathcal{K}$ that is simple, homotopic to γ, and no longer than γ. $\square$

In fact, the techniques used in the proof of this proposition yield an algorithm to compute a shortest cycle homotopic to a given boundary of a cross-metric cylinder or pair of pants; this will be discussed in more details in Section 6.

4.2 Optimal Paths on Pairs of Pants

Let $\mathcal{M}$ be a surface; let $\gamma : S^1 \to \mathcal{M}$ be a cycle. A path *wrapping around* γ is any path $p : [0,1] \to \mathcal{M}$ such that $p(t) = \gamma((at+b) \bmod 1)$ for some real numbers a and b. A boundary cycle ν of a cross-metric surface $\mathcal{M}$ is not regular with respect to $H_{\mathcal{M}}$; but the *length* of ν is naturally defined to be the sum of the lengths of the non-boundary edges of $H_{\mathcal{M}}$ adjacent to ν (counted with multiplicity). Similarly, the *length* of a path p wrapping around a boundary cycle ν is the sum of the lengths of the edges of $H_{\mathcal{M}}$ adjacent to ν and intersected by p, counted with multiplicity. In other words, the lengths of ν and p are the minimal lengths of slightly translated copies of these curves.

Proposition 4.4 *Let $\mathcal{K}$ be a cross-metric cylinder or a pair of pants; let ν be one boundary of $\mathcal{K}$. Assume ν is a shortest cycle among the simple cycles homotopic to ν. Then any path wrapping around ν is as short as possible in its homotopy class.*

The proof relies on the following lemma.

Lemma 4.5 *Let $\mathcal{P}$ be a cross-metric pair of pants. Let p be a path wrapping around a boundary ν of $\mathcal{P}$ and q be a path homotopic to p. Let r be a shortest path between two points, one on each of the two boundaries of $\mathcal{P}$ different from ν. Then there exists a path q', homotopic to and no longer than q, that does not cross r.*

PROOF. Analogous to Lemma 4.2 with crossing words instead of crossing word sets. $\square$

PROOF OF PROPOSITION 4.4. Let p be a path wrapping around ν and let q be homotopic to p in $\mathcal{K}$; we prove that $|p| \leq |q|$.

Assume first that $\mathcal{K}$ is a pair of pants. By Lemma 4.5, we may assume (up to changing q) that q does not cross a shortest path r between the two boundaries of $\mathcal{K}$ that are different from ν. On the cylinder $\mathcal{C}$ obtained by cutting $\mathcal{K}$ along r, the path q is homotopic to p. Indeed, on $\mathcal{C}$, q is homotopic to p concatenated with a power ν^k of ν; so, in $\mathcal{K}$, $q^{-1} \cdot p \cdot \nu^k$ is contractible, but also homotopic to ν^k; hence $k = 0$. It thus remains to prove the proposition when $\mathcal{K}$ is a cylinder.

So let $\mathcal{K}$ be a cylinder. The proof is by induction on the number $c(q)$ of self-crossings of q. If $c(q) = 0$, then q is homotopic to a simple subpath of ν, and the result follows since ν is a shortest simple cycle in its homotopy class. Assume that $c(q) > 0$ and that the result is true for all smaller values of $c(q)$. Let c_1 be a simple closed subpath of q, and let q_1 be equal to the path q where c_1 is removed; $c(q_1) \leq c(q) - 1$ and the path q_1 is homotopic to a path p_1 wrapping around ν.

The cycle c_1 is either contractible or freely homotopic to the boundaries of $\mathcal{K}$ [Eps66, Theorem 4.2]. In the former case, q is homotopic to q_1 and we conclude by the induction hypothesis. Otherwise, q is homotopic to either $q_1 \cdot \nu$ or $q_1 \cdot \nu^{-1}$, so that $|p| \leq |p_1| + |\nu|$; using the induction hypothesis, this cannot be greater than $|q_1| + |\nu|$, and, by the assumption on ν, this in turn cannot be greater than $|q_1| + |c_1| = |q|$. □

4.3 Optimal Cycles in (Doubled) Pants Decompositions

Proposition 4.6 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Let s be a (doubled) pants decomposition of $\mathcal{M}$ and let $\mathcal{K}$ be a component of the surface $\mathcal{M}$ cut along s ($\mathcal{K}$ is a cylinder or a pair of pants). Assume that a cycle γ inside $\mathcal{K}$ is homotopic in $\mathcal{M}$ to a cycle s_k. Then γ is homotopic* in $\mathcal{K}$ *to a boundary of $\mathcal{K}$.*

We will need the following lemma.

Lemma 4.7 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Let s be a (doubled) pants decomposition of $\mathcal{M}$ and let $\mathcal{K}$ be a component of the surface $\mathcal{M}$ cut along s. Any cycle inside $\mathcal{K}$ that is contractible in $\mathcal{M}$ is also contractible in $\mathcal{K}$.*

Proof. Consider the universal cover $(\widetilde{\mathcal{M}}, \pi)$ of $\mathcal{M}$; it is sufficient to prove that $\pi^{-1}(\mathcal{K})$ is simply connected. If it were not the case, there would exist a simple non-contractible cycle γ in $\pi^{-1}(\mathcal{K})$. Such a cycle bounds a disk D in $\widetilde{\mathcal{M}}$ that is not entirely contained in $\pi^{-1}(\mathcal{K})$. Therefore a lift $\widetilde{\nu}$ of a boundary ν of $\mathcal{K}$ is inside D. This is impossible since $\widetilde{\nu}$ contains infinitely many lifts of some point of $\mathcal{M}$ [Eps66, Lemma 4.3]. □

Proof of Proposition 4.6. Let $\widetilde{s}$ be the set of the geometric lifts of the curves in s. Let δ be a slightly translated copy of s_k (a simple cycle, disjoint from all the cycles in s, homotopic, in $\mathcal{M}$ cut along $s \setminus s_k$, to s_k). Let $p : [0,1] \to \mathcal{M}$ be the loop defined by $p(t) = \gamma(t \bmod 1)$; similarly, let $q : [0,1] \to \mathcal{M}$ be defined by $q(t) = \delta(t \bmod 1)$. There exists a path β joining $p(0)$ to $q(0)$ such that the path $r := \beta^{-1} \cdot p \cdot \beta \cdot q^{-1}$ is contractible in $\mathcal{M}$. Without loss of generality, assume that $\widetilde{s}/\widetilde{\beta}$ is irreducible for some, hence any, lift $\widetilde{\beta}$ of β. If this crossing word is empty, then r is contractible in $\mathcal{K}$ by Lemma 4.7, hence γ and δ are homotopic in $\mathcal{K}$; so are γ and s_k, and the proof is complete. Therefore, assume this crossing word is non-empty.

Let $\widetilde{r}$ be a lift of r. Since p and q do not cross s, $\widetilde{s}/\widetilde{r}$ is the concatenation of $\widetilde{s}/(\beta^0)^{-1}$ and $\widetilde{s}/\beta^1$, where β^0 is a lift of β and $\beta^1 = \tau_{\widetilde{\gamma}}(\beta^0)$. Because $\widetilde{s}/\beta^0$ and $\widetilde{s}/\beta^1$ are irreducible and $\widetilde{s}/\widetilde{r}$ can be elementarily reduced, the first geometric lifts of $\widetilde{s}$ crossed by β^0 and β^1 must be the same, say some geometric lift of s_j. Let β' be the beginning of β before its first crossing with s; we get that $\beta'^{-1} \cdot p \cdot \beta'$ is homotopic to a power of s_j in $\mathcal{M}$, hence also in $\mathcal{K}$ by Lemma 4.7. It is known [Eps66, Theorem 4.2] that the nth power of s_j is homotopic to no simple cycle if $|n| \geq 2$. Hence γ is homotopic in $\mathcal{K}$ to s_j or its reverse. □

5 Proof of Theorem 2.1

Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. In this section, we prove Theorem 2.1: if s is a doubled pants decomposition on $\mathcal{M}$, shortening the cycles of s using elementary steps provides, in a finite number of steps, a doubled pants decomposition in which each cycle is as short as possible in its homotopy class.

We fix some notations for this section. Let s be a doubled pants decomposition of $\mathcal{M}$. Fix $i, j \in [1, N]$. Let t_i be a shortest cycle homotopic to s_i and $\widetilde{t_i}$ be a lift of t_i. Let $r = f_j(s)$ be the doubled pants decomposition obtained from s by applying an elementary step to s_j. Consider $\mathcal{M}$ cut along $s \setminus s_j = r \setminus r_j$; the pair of pants of this surface that contains s_j and r_j is denoted by $\mathcal{P}_j$.

Let $\widetilde{s}$ and $\widetilde{r}$ be the sets of geometric lifts of the cycles in s and r, respectively. For $k \in [1, N]$, we choose an arbitrary indexation s_k^α, $\alpha \in \mathbb{N}$, of the geometric lifts of s_k. We define an indexation of the geometric lifts of r_k as follows. If $k \neq j$, let $r_k^\alpha = s_k^\alpha$ for each $\alpha \in \mathbb{N}$. Let us consider now the case $k = j$. Let δ be the twin of s_j (as defined at the end of Section 2.1); it is also the twin of r_j. The cycles s_j and δ bound a cylinder, which lifts to a set of disjoint infinite strips in $\widetilde{\mathcal{M}}$. This provides a correspondence between the geometric lifts of s_j and the geometric lifts of δ. Similarly, the fact that δ and r_j bound a cylinder provides a correspondence between the geometric lifts of δ and the geometric lifts of r_j. By composition, we obtain a correspondence between the geometric lifts of s_j and those of r_j; for each geometric lift s_j^α, we define r_j^α as the geometric lift of r_j in correspondence with s_j^α.

We will consider crossing words between $\widetilde{s}$ (or $\widetilde{r}$) and $\widetilde{t_i}$. Henceforth, the words in the symbols a and $\overline{a}$, where $a \in \widetilde{s}$ (or $a \in \widetilde{r}$) will be written differently as above: we only write the subscripts and superscripts corresponding to the geometric lifts (for example, we shall write ${}^{3}_{1}\,\overline{{}^{7}_{5}}\,\overline{{}^{4}_{2}}$ instead of $s_1^3\,\overline{s_5^7}\,\overline{s_2^4}$). This allows to say, for example, that $[\widetilde{r}/\widetilde{t_i}] = [\widetilde{s}/\widetilde{t_i}]$ if t_i does not cross r_j nor s_j. Similarly, we use the expression *j-reduction* instead of s_j-reduction or r_j-reduction.

If W is a $\widetilde{t_i}$-word set, recall that $g_j^{\widetilde{t_i}}(W)$ is the unique $\widetilde{t_i}$-word set to which W j-reduces.

Proposition 5.1 $g_j^{\widetilde{t_i}}([\widetilde{r}/\widetilde{t_i}]) = g_j^{\widetilde{t_i}}([\widetilde{s}/\widetilde{t_i}])$.

We denote by $\widetilde{s}_j$ the set of the geometric lifts of s_j, with the indexation inherited from $\widetilde{s}$. We define $\widetilde{r}_j$ analogously. We will use the following lemma.

Lemma 5.2 *Let p be an arc in $\mathcal{P}_j$, and let $\widetilde{p}$ be a lift of p. Then $\widetilde{s}_j/\widetilde{p}$ and $\widetilde{r}_j/\widetilde{p}$ reduce to the same irreducible word.*

Proof. We first assume s_j and r_j are disjoint; they bound a cylinder $\mathcal{C}$ inside $\mathcal{P}_j$. The lifts of $\mathcal{C}$ in the universal cover of $\mathcal{M}$ are pairwise disjoint infinite strips bounded by r_j^α and s_j^α, for all α (by the choice of the indexation of the geometric lifts of s_j and r_j). Furthermore,

$\widetilde{p}$ has its endpoints outside these strips. Let us split $\widetilde{p}$ into subpaths $\widetilde{p}_i$, $i = 1, \ldots, n$, each entering in exactly one strip, and exactly once in this strip, and so that their endpoints are outside the strips. Clearly, $\widetilde{s}_j/\widetilde{p}_i$ and $\widetilde{r}_j/\widetilde{p}_i$ reduce to the same irreducible word (there are two cases according to whether $\widetilde{p}_i$ enters and exits the strip through the same boundary or not); the result follows.

For the general case, let γ be a slightly translated copy of the twin of s_j: γ is a simple cycle homotopic in $\mathcal{P}_j$ to r_j and s_j, and does not cross r_j nor s_j. We index the geometric lifts of γ so that γ^α and r_j^α (or s_j^α) bound an infinite strip. Applying the reasoning of the above paragraph to s_j and γ, and then to γ and r_j, we get the result. $\square$

Proof of Proposition 5.1. Assume first that t_i is contained in $\mathcal{P}_j$. By Proposition 3.7, we have $g^{\widetilde{t}_i}([\widetilde{r}/\widetilde{t}_i]) = \{\varepsilon\} = g^{\widetilde{t}_i}([\widetilde{s}/\widetilde{t}_i])$. But this also equals $g_j^{\widetilde{t}_i}([\widetilde{r}/\widetilde{t}_i])$ and $g_j^{\widetilde{t}_i}([\widetilde{s}/\widetilde{t}_i])$, and this concludes the proof. If t_i is not entirely contained in $\mathcal{P}_j$, then let t'_i be a maximal subpath of t_i that is inside $\mathcal{P}_j$, and $\widetilde{t}'_i$ be a lift of t'_i; it is sufficient to prove that $\widetilde{r}_j/\widetilde{t}'_i$ and $\widetilde{s}_j/\widetilde{t}'_i$ reduce to the same irreducible word; but this follows from Lemma 5.2. $\square$

The following key proposition states that an elementary step f_j on the doubled pants decomposition s corresponds to a reduction $g_j^{\widetilde{t}_i}$ on the crossing word set $[\widetilde{s}/\widetilde{t}_i]$. This is not entirely true, because we may need to replace t_i by a cycle t'_i that has exactly the same properties as t_i (if there exist several shortest homotopic cycles).

Proposition 5.3 *There exists a cycle t'_i, homotopic to and no longer than t_i, and a lift $\widetilde{t}'_i$ of t'_i, such that $\tau_{\widetilde{t}_i} = \tau_{\widetilde{t}'_i}$ and $[\widetilde{r}/\widetilde{t}'_i] = g_j^{\widetilde{t}_i}([\widetilde{s}/\widetilde{t}_i])$.*

The fact that $\tau_{\widetilde{t}_i} = \tau_{\widetilde{t}'_i}$ implies in particular that $g_k^{\widetilde{t}_i} = g_k^{\widetilde{t}'_i}$ for each k: although we replace $\widetilde{t}_i$ by $\widetilde{t}'_i$, the operations $g_k^{\widetilde{t}_i}$ remain the same.

Proof. By Proposition 5.1, $[\widetilde{r}/\widetilde{t}_i]$ j-reduces to $g_j^{\widetilde{t}_i}([\widetilde{s}/\widetilde{t}_i])$. If $[\widetilde{r}/\widetilde{t}_i]$ is j-irreducible, there is nothing to show. Otherwise, an elementary j-reduction is possible on $[\widetilde{r}/\widetilde{t}_i]$. We apply the uncrossing operation to the subpath p of $\widetilde{t}_i$ corresponding to this j-reduction. By Proposition 4.4, applied to the component $\mathcal{K}$ of $\mathcal{M} \setminus r$ containing the projection of the subpath p, we obtain a lift $\widetilde{t}'_i$ of a cycle t'_i that is homotopic to and no longer than t_i. Clearly, $\tau_{\widetilde{t}_i} = \tau_{\widetilde{t}'_i}$. Furthermore, $[\widetilde{r}/\widetilde{t}'_i]$ results from $[\widetilde{r}/\widetilde{t}_i]$ by this elementary j-reduction. By induction, we obtain the desired t'_i. $\square$

Proposition 5.4 *Let $k \in [1, N]$. Assume that t_i is disjoint from s and that t_i and s_k are homotopic in the cylinder or pair of pants of the surface $\mathcal{M}$ cut along s containing t_i. Then, there exists a cycle t'_i, homotopic to and no longer than t_i, disjoint from $r = f_j(s)$, and homotopic to r_k in the cylinder or pair of pants of $\mathcal{M}$ cut along r containing t'_i.*

Proof. Let $\mathcal{K}$ be the cylinder or pair of pants of the surface $\mathcal{M}$ cut along s containing t_i. The proof is trivial if s_j is not a boundary of $\mathcal{K}$. If s_j is homotopic to t_i in $\mathcal{K}$, then either $s_j = s_k$, or $\mathcal{K}$ is the cylinder bounded by s_j and s_k; in both cases it is easy to conclude. Indeed, t_i can be chosen to be simple by Proposition 4.1; it follows that $|r_j| = |t_i|$, and we can take for t_i' a slightly translated copy of r_j.

There remains the case where s_j is not homotopic to t_i in $\mathcal{K}$ (and thus $\mathcal{K}$ is a pair of pants). $\mathcal{P}_j$ contains t_i; one boundary of $\mathcal{P}_j$ is γ, the twin of r_j, and another one is $r_k = s_k$. r_j and γ bound a cylinder $\mathcal{C}$ in $\mathcal{P}_j$, and r_j is optimal in $\mathcal{C}$. Then, using Proposition 4.4, any component of t_i in $\mathcal{C}$ can be swapped into the complementary part in $\mathcal{P}_j$, thus removing the crossings with r_j. □

We now conclude the proof of our main theorem.

Proof of Theorem 2.1. In the beginning of this section, we explained how, given an arbitrary indexation of the geometric lifts of a doubled pants decomposition s, we deduce an indexation of the geometric lifts of $r = f_j(s)$. Now, starting with an arbitrary indexation of the geometric lifts of s^0, we can proceed recursively to define an indexation of the geometric lifts of all the doubled pants decompositions occurring during the algorithm (deduced from s^0 by successive applications of elementary steps). Let $\widetilde{s}^n$ denote the geometric lifts of the curves in s^n with this particular indexation.

Let $i \in [1, N]$ and let $\widetilde{t}_i^0$ be a lift of a shortest cycle t_i^0 homotopic to s_i^0 in $\mathcal{M}$. By Proposition 3.7, $[\widetilde{s}^0/\widetilde{t}_i^0]$ reduces to $\{\varepsilon\}$. By Proposition 5.3, we can construct a sequence $(\widetilde{t}_i^n)_{n\in\mathbb{N}}$ of lifts of shortest homotopic cycles such that the length of $[\widetilde{s}^n/\widetilde{t}_i^n]$ strictly decreases until it becomes empty at some stage n. By Proposition 4.6, t_i^n and a cycle s_k^n are homotopic in the cylinder or pair of pants of the surface $\mathcal{M}$ cut along s^n containing t_i^n. By $k-1$ applications of Proposition 5.4, and then using Proposition 4.1, $|s_k^{n+1}| = |t_i^n|$. The cycle s_k^{n+1} is either s_i^{n+1} or its twin; in the latter case, since s_i^{n+1} and s_k^{n+1} bound a cylinder, $|s_i^{n+2}| = |s_k^{n+2}| = |t_i^n|$. From this discussion, it follows that the length of $(s_i^n)_{n\in\mathbb{N}}$ becomes stationary and that, when it is the case, each cycle is as short as possible in its homotopy class. It remains to prove that all lengths remain unchanged once s^n and s^{n+1} have the same total length. Assume s and $s' = f(s)$ have the same length, and let $i \in [1, N]$; we shall prove that s_i has the same length as t_i (a shortest cycle homotopic to s_i).

$[\widetilde{s}/\widetilde{t}_i]$ reduces to the empty word set; assume that an elementary j-reduction is possible. Let $\overline{t}_i$ and $\overline{s}_j$ be the associated subpaths of $\widetilde{t}_i$ and of the lift of s_j. We will prove that both subpaths have the same length. It will follow that we can modify t_i without changing its length nor its homotopy class to proceed to the j-reduction in $[\widetilde{s}/\widetilde{t}_i]$; hence by induction we will be able to assume that $[\widetilde{s}/\widetilde{t}_i] = \{\varepsilon\}$.

If $j \neq 1$, only lifts of the first cycle in $f_1(s)$ appear in the word $\widetilde{f_1(s)}/\overline{t}_i$; by Corollary 3.4, this word is parenthesized. By Proposition 4.4, we can iteratively remove all the crossings between $\overline{t}_i$ and the lifts of $f_1(s)$. The path $\overline{t}_i$ is replaced this way by a no longer, homotopic path $\overline{t}_i'$ that does not cross the lifts of $f_1(s)$. Iterating the process, we get the existence of a path $\overline{t}''_i$, no longer than and homotopic to $\overline{t}_i$, crossing no lift of any cycle in $s'' := f_{j-1} \circ \ldots \circ f_1(s)$. Furthermore, $s''_j = s_j$ is as short as possible in its homotopy class in the cylinder and in the pair of pants it bounds, because s'' has the same length as $f_j(s'')$. It follows, by Proposition 4.4, that $\overline{s}_j$ cannot be longer than $\overline{t}''_i$. Hence $|\overline{s}_j| = |\overline{t}''_i| = |\overline{t}_i|$.

We can thus assume (up to a change of t_i) that $[\widetilde{s}/\widetilde{t_i}] = \varepsilon$. By Proposition 4.6, t_i and a cycle s_k are homotopic in the cylinder or pair of pants of the surface $\mathcal{M}$ cut along s containing t_i. By $k-1$ applications of Proposition 5.4, and then by Proposition 4.1, we may assume that t_i and s_k bound a cylinder whose interior is disjoint from the cycles of $f_{k-1} \circ \ldots \circ f_1(s)$. This implies $|s_k| = |t_i|$, which finishes the proof if $k = i$. If $k \neq i$, s_k is the twin of s_i; s_k and s_i bound a cylinder whose interior is disjoint from s; then we must have $|s_k| = |s_i|$ (otherwise, the length of the longest cycle would decrease by applying f to s). This concludes the proof. □

6 Computational Issues

This section explains how Theorem 2.1 can be turned into a practical algorithm for optimizing (doubled) pants decompositions. Recall that, as explained in Section 2.2, we perform all our operations using merely the combinatorial properties of the arrangement of the graph $H_{\mathcal{M}}$ and of the curves.

6.1 Computation of an Elementary Step

The algorithm is a succession of elementary steps; each of them amounts to computing, in a cross-metric pair of pants, a shortest simple cycle homotopic to one of its boundaries. The following proposition explains how to perform these steps.

Proposition 6.1 *Let $\mathcal{P}$ be a cross-metric pair of pants of complexity n, and let γ be a boundary cycle of $\mathcal{P}$. We can compute a shortest cycle homotopic to γ in $\mathcal{P}$ in $O(n \log n)$ time, with the additional property that this cycle is simple.*

We will rely on the following lemma.

Lemma 6.2 *Let $\mathcal{C}$ be a cross-metric cylinder of complexity n. A shortest cycle homotopic to the boundaries of $\mathcal{C}$, with the additional property that it is a simple cycle, can be computed in $O(n \log n)$ time.*

Proof. Let γ be a shortest cycle homotopic to the boundaries of $\mathcal{C}$. By proposition 4.1 we may assume γ is simple. We claim that γ crosses each edge of $H_{\mathcal{C}}$ at most once. Suppose on the contrary that γ crosses some edge e at least twice and consider two consecutive crossings a and b of γ along e. Then, using paths joining a and b along e we may split γ into two simple cycles shorter than γ. One of these two cycles must be homotopic to γ, a contradiction.

So the problem amounts to computing a minimal cut in $H_{\mathcal{C}}$ separating the two boundaries of $\mathcal{C}$ [Rei83, Propositions 1 and 2]. Since $H_{\mathcal{C}}$ is a planar graph, this can be done in $O(n \log n)$ time using an algorithm by [Fre87, Theorem 7]. $\square$

Proof of Proposition 6.1. By the proof of Proposition 4.1, this can be done as follows:

- Compute a shortest path p between the two boundaries of $\mathcal{P}$ that are not homotopic to γ in $\mathcal{P}$;
- compute a shortest simple cycle homotopic to the boundaries of the cylinder $\mathcal{C}$ obtained by cutting $\mathcal{P}$ along p.

Recall that, on any cross-metric surface $\mathcal{M}$, we can compute shortest paths between two points by considering any shortest-path algorithm in the dual of the graph $H_{\mathcal{M}}$; we can also (temporarily) cut $\mathcal{M}$ along a specified set of curves by declaring that the corresponding edges cannot be crossed. Since both operations described above boil down to the computation of shortest paths and cutting along some curves, we can achieve this task in cross-metric surfaces.

More precisely, the first step can be done using Dijkstra's algorithm in $O(n \log n)$ time, where n is the complexity of $\mathcal{P}$. Lemma 6.2 states that the second step can also be done in $O(n \log n)$ time. This concludes the proof. $\square$

6.2 Complexity Analysis

We give an analysis of the running-time of our algorithm.

Theorem 6.3 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Assume $\mathcal{M}$ has complexity n, genus g, and number of boundaries b. Let α be the ratio between the largest and smallest weight of $\mathcal{M}$. Let $s = (s_1, \dots, s_N)$ be a doubled pants decomposition of $\mathcal{M}$ where each cycle has complexity $O(m)$.*

We can compute a doubled pants decomposition $u = (u_1, \dots, u_N)$ of $\mathcal{M}$ such that, for each i, u_i is a shortest cycle among all (not necessarily simple) cycles homotopic to s_i, in $O((g+b)^2 \alpha m^2 (\alpha m + n) \log(\alpha m + n))$ time.

We note that the time complexity of this algorithm is polynomial in its input and in α. In particular, if the weights are uniform, then this algorithm has a running-time that is polynomial in the size of the input.[4]

[4] Very recently, [CdVE06] prove, using so-called *tight octagonal decompositions* or *tight systems of arcs* of the surface, that this algorithm has actually polynomial running-time in the size of its input, independently from α. As a consequence, the algorithm for computing shortest homotopic cycles (Theorem 7.2) also has polynomial running-time.

PROOF. Let $i \in [1, N]$. Let t_i be a shortest cycle homotopic to s_i. We can choose t_i so that, inside each face f of $H_{\mathcal{M}}$, it crosses each maximal subpath of the cycles in s entering f at most once; hence the number of crossings of t_i with s is at most the complexity of t_i times the complexity of s. Since t_i is no longer than s_i, its complexity cannot be larger than α times the complexity of s_i; so the complexity of t_i is $O(\alpha m)$. Thus the crossing word set $[\widetilde{s}/\widetilde{t_i}]$ has length at most $O((g+b)\alpha m^2)$ since there are $O(g+b)$ cycles. The number of elementary steps is $O(g+b)$ times the length of the longest crossing word set $[\widetilde{s}/\widetilde{t_i}]$ over all i; this is thus $O((g+b)^2\alpha m^2)$.

Since the length of the doubled pants decomposition does not increase during the course of the algorithm, the complexity of each of its cycles is bounded, at each step, by $O(\alpha m)$. It follows that $\mathcal{P}_j$, bounded by three such cycles, has complexity $O(\alpha m + n)$. The cost of each elementary step is thus $O((\alpha m + n)\log(\alpha m + n))$ by Proposition 6.1. The result follows. □

7 Computing a Shortest Homotopic Cycle

In this section, we focus on the following problem: given a simple cycle s_1 on a cross-metric surface $\mathcal{M}$, compute a simple cycle among the shortest cycles in its homotopy class.

The following proposition will be proved by two different methods in Sections 7.1 and 7.2. The analysis of the second algorithm gives an extra $O(n \log n)$ term in the time complexity, but the method has a more geometric flavor and might be interesting for design purposes.

Proposition 7.1 *Let $\mathcal{M}$ be a cross-metric surface of complexity n, genus g, and b boundaries that is not a sphere, disk, cylinder, or torus. We can compute a pants decomposition of $\mathcal{M}$ where each cycle has complexity $O(n)$, in $O((g+b)n)$ time.*

Admitting temporarily this result, we obtain:

Theorem 7.2 *Let $\mathcal{M}$ be a cross-metric surface that is not a sphere, disk, cylinder, or torus. Assume $\mathcal{M}$ has complexity n, genus g, and number of boundaries b. Let α be the ratio between the largest and smallest weight of $\mathcal{M}$. Let s_1 be a simple cycle on $\mathcal{M}$ with complexity k. We can compute a shortest cycle homotopic to s_1, with the additional property that it is simple, in $O((g+b)^2\alpha^2(k+n)^3 \log(\alpha(k+n)))$ time.*

PROOF. We cut $\mathcal{M}$ along s_1 in $O(k)$ time, obtaining one surface $\mathcal{M}_1$ or two surfaces $\mathcal{M}_1$ and $\mathcal{M}_2$. We can test in $O(n+k)$ time whether one of these surfaces is a disk, which happens if and only if s_1 is contractible; in this case the problem is trivial. We compute pants decompositions of the $\mathcal{M}_i$'s that are not cylinders in time $O((g+b)(k+n))$ using Proposition 7.1; each cycle has complexity $O(k+n)$.

On $\mathcal{M}$, the union of s_1 and of these cycles is a decomposition of $\mathcal{M}$ into cylinders and pairs of pants. Appending slightly translated copies of some cycles, we obtain a doubled pants decomposition of $\mathcal{M}$. We can then apply our iterative algorithm for shortening a doubled pants decomposition; by Theorem 2.1, the cycle of the resulting doubled pants decomposition that corresponds to s_1 has the desired properties; by Theorem 6.3, this process has the indicated running-time. $\square$

Again, this result also holds if $\mathcal{M}$ is a cylinder or a torus; we omit the details.

We define the *edge multiplicity* of a cycle on a cross-metric surface $\mathcal{M}$ as the maximum, over all edges e of $H_{\mathcal{M}}$, of the number of crossings of this cycle with e.

7.1 Pants Decomposition I

PROOF OF PROPOSITION 7.1. We say that a set of cycles s is *good* if the cycles are simple and pairwise disjoint and if, for each connected component $\mathcal{M}'$ of the surface $\mathcal{M}$ cut along the cycles in s:

1. $\mathcal{M}'$ has at least one boundary, and is not a disk nor a cylinder;
2. if $\mathcal{M}'$ is a pair of pants, each edge e of $H_{\mathcal{M}}$ not on the boundary of $\mathcal{M}$ crosses the boundary of $\mathcal{M}'$ at most four times;
3. if $\mathcal{M}'$ is not a pair of pants, each edge e of $H_{\mathcal{M}}$ not on the boundary of $\mathcal{M}$ crosses the boundary of $\mathcal{M}'$ at most twice.

If $\mathcal{M}$ has at least one boundary, $s = \emptyset$ is a good set of cycles. If $\mathcal{M}$ has no boundary, any non-contractible cycle γ with edge multiplicity one constitutes itself a good set of cycles. We can compute such a γ in $O(n)$ time, using the technique by [EHP04, Corollary 5.3] with unit weights and breadth-first search instead of Dijkstra's algorithm. Or, alternately, such a γ can be obtained by considering a spanning tree of $H_{\mathcal{M}}$; any simple cycle in the dual graph restricted to the duals of the non-tree edges will do.

By Condition 1, every good set of cycles can be extended to a pants decomposition of $\mathcal{M}$. Given a good set of cycles s that is not a pants decomposition, we will see how to append to s one or two cycles, so that the resulting set of cycles is still good. We finally obtain a pants decomposition made of $O(g + b)$ cycles, each of multiplicity at most four, hence each of complexity $O(n)$. The assertion on the running-time is easy and follows from Conditions 2 and 3.

Let s be a good set of cycles that is not a pants decomposition. We distinguish between two cases.

First case. Assume that there exists a connected component $\mathcal{M}'$ of the surface $\mathcal{M}$ cut along s such that its number of boundaries b' and its genus g' satisfy $b' \geq 4$ and/or ($g' \geq 1$ and $b' \geq 2$). We will append to s a cycle "merging" two boundaries of $\mathcal{M}'$ (Figure 7): the new cycle will thus split $\mathcal{M}'$ into a pair of pants and a surface with $b' - 1$ boundaries and of genus g'.

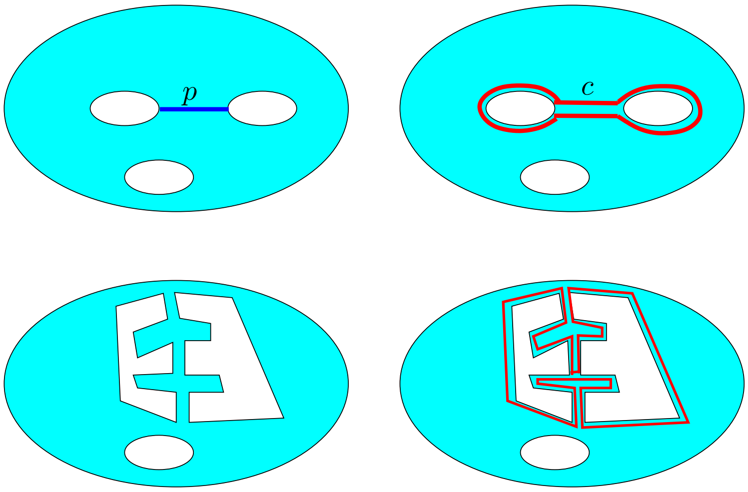

Figure 7: In this figure, curves that run along close together are assumed to be crossed by the same edges of $H_{\mathcal{M}'}$. Top: cutting $\mathcal{M}'$ with a path p joining two boundaries of $\mathcal{M}'$. Bottom: several edges of $H_{\mathcal{M}'}$ are crossed by two boundaries of $\mathcal{M}'$; in that case, p has length zero.

Choose a boundary B_1 of $\mathcal{M}'$; let B_2 be the union of the other boundaries of $\mathcal{M}'$. Using breadth-first search, we compute a shortest path p (for uniform weights) with edge multiplicity one on $\mathcal{M}'$ between B_1 and B_2. We then "merge" the two boundaries of $\mathcal{M}'$ joined by p with a cycle c that goes once around each of the two boundaries and twice along p. Appending c to s, we obtain a set of cycles s'.

We claim that s' is good. First note that p crosses no edge of $H_{\mathcal{M}'}$ incident to a boundary of $\mathcal{M}'$, for otherwise it would not be a shortest path between B_1 and B_2. So the edges of $H_{\mathcal{M}}$ crossed by p and those crossed by a boundary of $\mathcal{M}'$ are disjoint. It also follows that p crosses each edge of $H_{\mathcal{M}}$ at most once.

The cycle c separates $\mathcal{M}'$ into two connected components:

- a pair of pants, whose boundary crosses the following two sets of edges of $H_{\mathcal{M}}$; these sets are disjoint by the above remark:
  - twice each edge of $H_{\mathcal{M}}$ crossed by b, for each boundary b of $\mathcal{M}'$ containing an endpoint of p. By the induction hypothesis, an edge of $H_{\mathcal{M}}$ crosses the boundaries of $\mathcal{M}'$ at most twice, so each edge of $H_{\mathcal{M}}$ crosses the boundary of this pair of pants at most four times;
  - twice each edge of $H_{\mathcal{M}}$ crossed by p (recall that p crosses a given edge of $H_{\mathcal{M}}$ at most once);
- the complementary part of this pair of pants. The edges of $H_{\mathcal{M}}$ crossed by its

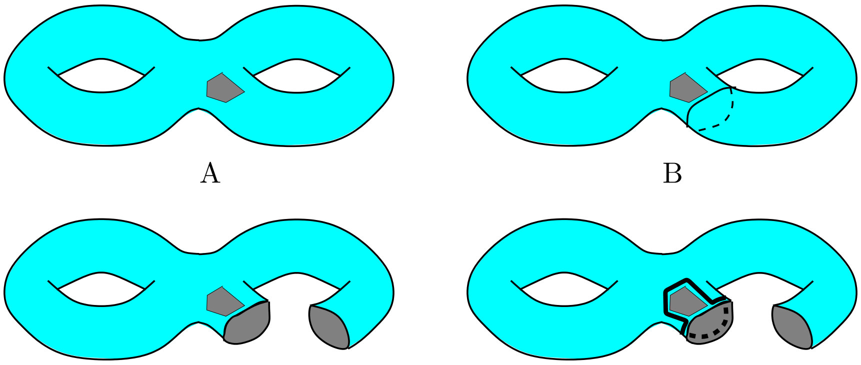

Figure 8: Cutting a double-torus with one boundary $\mathcal{M}'$ along an essential cycle. A: the surface $\mathcal{M}'$. B: Creation of an essential cycle (here, non-separating) adjacent to the boundary. C: Cutting the surface along this cycle. D: Creation of a new cycle enclosing the two adjacent boundaries.

boundaries consist of the following two disjoint sets of edges:

- those that are crossed by the boundaries of $\mathcal{M}'$: by the induction hypothesis, each edge of $H_{\mathcal{M}}$ crosses these boundaries at most twice;
- the edges of $H_{\mathcal{M}}$ crossed by p; each such edge is crossed twice by the boundary of this surface.

These remarks imply that s is good.

Second case. A connected component $\mathcal{M}'$ of $\mathcal{M}$ cut along s has exactly one boundary, denoted by b, and has non-zero genus. The idea is to cut $\mathcal{M}'$ along an essential cycle and then along another cycle to enforce that the set of cycles is good (Figure 8).

Let $\overline{\mathcal{M}}'$ be the cross-metric surface obtained from $\mathcal{M}'$ by closing the boundary b with an extra face f and assigning unit weights to all edges. Using the technique by [EHP04, Lemma 5.2], with a breadth-first search instead of Dijkstra's algorithm, we obtain, in time linear in the complexity of $\overline{\mathcal{M}}'$, a shortest non-contractible loop ℓ based at a point inside f that crosses each edge of $\overline{\mathcal{M}}'$ at most twice and crosses b exactly twice. Observe that ℓ does not cross any edge of $H_{\overline{\mathcal{M}}'}$ incident to b: otherwise, one could built two loops ℓ_1 and ℓ_2 based inside f, whose concatenation is homotopic to ℓ and shorter than ℓ; then at least one of ℓ_1 or ℓ_2 would be non-contractible and shorter than ℓ. We eventually form a cycle c in $\mathcal{M}'$ from ℓ by replacing the part of ℓ inside f by a simple path that "runs along" b outside f. See Figure 8B. The cycle c is essential in $\mathcal{M}'$, because ℓ is non-contractible in $\overline{\mathcal{M}}'$. We append c to s.

The set of cycles s may fail to be good because of the edges of $H_{\mathcal{M}}$ crossed by both b and c (Figure 8C). But we can easily remedy this problem by enclosing b and c with a new cycle d (Figure 8D) that separates the surface into two surfaces:

- a pair of pants, whose boundary crosses each edge of $H_{\mathcal{M}}$ at most four times (as in the first case),
- and a surface with one or two boundaries, according to whether c is non-separating or not. Assume c is non-separating (the case where it is separating is simpler). The surface is made of two boundaries: a copy of d and a copy of c. Recall that b crosses each edge of $H_{\mathcal{M}}$ at most twice by the induction hypothesis; so on the present surface, each edge of $H_{\mathcal{M}}$ is crossed at most twice by its boundaries as well.

It follows that the current set of cycles is good. $\square$

7.2 Pants Decomposition II

Given a cross-metric surface $\mathcal{M}$ of complexity n, we now outline a second method to compute a pants decomposition of $\mathcal{M}$ where each cycle has edge multiplicity bounded by a constant. The idea is to mimic the procedure of [Hat00] in the differentiable case to compute a pants decomposition from the level sets of a piecewise-linear (PL) function on $\mathcal{M}$. We assume familiarity with the basic notion of simplicial complex [Hat02, Section 2.1]. Again, we emphasize that we prove here Proposition 7.1 with an additional $O(n \log n)$ term in the time complexity.

Proof of Proposition 7.1, with a weaker complexity. We first identify each boundary component of $\mathcal{M}$ to a single vertex by contracting the boundary edges. We obtain a surface $\mathcal{M}'$ without boundary. We call *β-vertices* the vertices of $\mathcal{M}'$ corresponding to the boundaries of $\mathcal{M}$.

We then turn $\mathcal{M}'$ into a simplicial complex by applying two consecutives "barycentric subdivisions". For this, we view $\mathcal{M}'$ as a gluing of k-gons, $k \geq 1$, where a k-gon is a topological disk bounded by k simple curves, called sides, and k vertices. Because of 1-gons and because of the identification of sides by the gluing, there might be multi-incidences between vertices, edges, and faces of $\mathcal{M}'$. A barycentric subdivision is obtained by subdividing each face before the gluing. This subdivision adds a vertex in the interior of each face (a face-vertex) and splits each side of a face with an interior vertex; it then joins each face-vertex to all the vertices on the boundary of its face with new edges. Call sd $\mathcal{M}'$ the surface obtained from $\mathcal{M}'$ after a barycentric subdivision. The faces of sd $\mathcal{M}'$ are 3-gons and sd $\mathcal{M}'$ has no multi-incidence; but two distinct edges or two distinct faces of sd $\mathcal{M}'$ may still share the same set of vertices. A second barycentric subdivision solves this problem and yields a surface $\mathrm{sd}^2\mathcal{M}'$ that is an abstract simplicial complex. Note that two β-vertices of $\mathcal{M}'$ cannot be adjacent on $\mathrm{sd}^2\mathcal{M}'$.

The next step is to construct a PL function on $\mathrm{sd}^2\mathcal{M}'$ with the requirement that the β-vertices are local minima and that the function has only simple singularities (see Tarasov and Vyalyi [TV98] or [CSA03] for the relevant definitions). We assign a different scalar value to each vertex of $\mathrm{sd}^2\mathcal{M}'$, giving smaller values to β-vertices. Extending this scalar field by linearity, we get a PL function which may have non-simple singularities. But, up to refining $\mathrm{sd}^2\mathcal{M}'$ at most twice, we obtain a surface $\mathcal{N}$ and a PL function f on $\mathcal{N}$ with simple singularities only [TV98, CSA03].[5]

Following [Hat00], we finally construct a pants decomposition of $\mathcal{N}$ (punctured at the β-vertices) with the help of the contour graph $\Gamma(f)$ of f. Because f has simple singularities, $\Gamma(f)$ has vertices of degree one or three only. We recursively remove the degree one vertices of $\Gamma(f)$ except those corresponding to β-vertices. Next, we merge the arcs of the resulting graph now sharing a degree two vertex. This way, we get a cubic graph with dangling arcs "incident" to β-vertices. The pants decomposition is obtained by picking a point in the interior of each non-dangling arc and considering the associated contours. Since a contour can cross an edge of $\mathcal{N}$ at most once, and since the edges of $\mathcal{M}'$ have been subdivided $O(1)$ times, we eventually get the result by considering the contours on the original surface $\mathcal{M}$ instead of $\mathcal{M}'$.

Since we only use a finite number of barycentric or similar subdivisions, the surface $\mathcal{N}$ has complexity $O(n)$ and can be constructed in time proportional to its complexity. It is known [CMEH+04] that $\Gamma(f)$ can be constructed in $O(n \log n)$ time. Each of the $O(g+b)$ contour cycles can be obtained by a simple traversal in time proportional to its size $O(n)$. This gives a total complexity of $O((g+b)n + n\log n)$ for the construction of the pants decomposition. □

Acknowledgements

We would like to thank Jeff Erickson for his numerous suggestions, which allowed us to significantly improve the presentation of the paper.

References

[Bes03] S. Bespamyatnikh. Computing homotopic shortest paths in the plane. *Journal of Algorithms*, 49(2):284–303, 2003.

[Bus92] P. Buser. *Geometry and spectra of compact Riemann surfaces*, volume 106 of *Progress in Mathematics*. Birkhäuser, Boston-Basel-New York, 1992.

[CdVE06] É. Colin de Verdière and J. Erickson. Tightening non-simple paths and cycles on surfaces. In *Proceedings of the 17th Annual Symposium on Discrete Algorithms*, pages 192–201, Miami, FL, 2006. ACM-SIAM.

[5] The authors of the aforementioned papers [TV98, CSA03] make use of a global linear function over the surface, assuming the surface is PL embedded in some $\mathbb{R}^d$. A closer look at their perturbation scheme shows that such an embedding is actually unnecessary.

[CdVL05] É. Colin de Verdière and F. Lazarus. Optimal system of loops on an orientable surface. *Discrete & Computational Geometry*, 33(3):507–534, 2005.

[CM07] S. Cabello and B. Mohar. Finding shortest non-separating and non-contractible cycles for topologically embedded graphs. *Discrete & Computational Geometry*, 37(2):213–235, 2007.

[CMEH$^+$04] K. Cole-McLaughlin, H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci. Loops in Reeb graph of 2-manifolds. *Discrete & Computational Geometry*, 32(2):231–244, 2004.

[CSA03] Hamish Carr, Jack Snoeyink, and Ulrike Axen. Computing contour trees in all dimensions. *Computational Geometry: Theory and Applications*, 24:75–94, 2003.

[DG99] T. K. Dey and S. Guha. Transforming curves on surfaces. *Journal of Computer and System Sciences*, 58:297–325, 1999.

[dGS97] M. de Graaf and A. Schrijver. Making curves minimally crossing by Reidemeister moves. *Journal of Combinatorial Theory, Series B*, 70(1):134–156, 1997.

[EHP04] J. Erickson and S. Har-Peled. Optimally cutting a surface into a disk. *Discrete & Computational Geometry*, 31(1):37–59, 2004.

[EKL06] A. Efrat, S. G. Kobourov, and A. Lubiw. Computing homotopic shortest paths efficiently. *Computational Geometry: Theory and Applications*, 35(3):162 – 172, 2006.

[Eps66] D. B. A. Epstein. Curves on 2-manifolds and isotopies. *Acta Mathematica*, 115:83–107, 1966.

[EW05] J. Erickson and K. Whittlesey. Greedy optimal homotopy and homology generators. In *Proceedings of the 16th Annual Symposium on Discrete Algorithms*, pages 1038–1046, Vancouver, BC, 2005. ACM-SIAM.

[FHS82] M. Freedman, J. Hass, and P. Scott. Closed geodesics on surfaces. *Bulletin of the London Mathematical Society*, 14:385–391, 1982.

[Fre87] G. N. Frederickson. Fast algorithms for shortest paths in planar graphs, with applications. *SIAM Journal on Computing*, 16(6):1004–1022, 1987.

[Hat00] A. Hatcher. Pants decompositions of surfaces. `http://www.math.cornell.edu/~hatcher/Papers/pantsdecomp.pdf`, 2000.

[Hat02] A. Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge - New York, 2002. Available at `http://www.math.cornell.edu/~hatcher/`.

[HS85] J. Hass and P. Scott. Intersections of curves on surfaces. *Israel Journal of Mathematics*, 51(1–2):90–120, 1985.

[HS94a] J. Hass and P. Scott. Shortening curves on surfaces. *Topology*, 33(1):25–43, 1994.

[HS94b] J. Hershberger and J. Snoeyink. Computing minimum length paths of a given homotopy class. *Computational Geometry: Theory and Applications*, 4:63–98, 1994.

[Kut06] M. Kutz. Computing shortest non-trivial cycles on orientable surfaces of bounded genus in almost linear time. In *Proceedings of the 22nd Annual Symposium on Computational Geometry*, pages 430–438, Sedona, AZ, 2006. ACM.

[LPVV01] 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 Symposium on Computational Geometry*, pages 80–89, Medford, MA, 2001. ACM.

[Mas77] W. S. Massey. *Algebraic Topology: An Introduction*, volume 56 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1977.

[Mit00] J. S. B. Mitchell. Geometric shortest paths and network optimization. In Jörg-Rüdiger Sack and Jorge Urrutia, editors, *Handbook of computational geometry*, pages 633–701. Elsevier, Amsterdam, 2000.

[MT01] B. Mohar and C. Thomassen. *Graphs on surfaces.* Johns Hopkins Studies in the Mathematical Sciences. John Hopkins University Press, Baltimore, MD, 2001.

[Rei83] J. H. Reif. Minimum $s-t$ cut of a planar undirected network in $O(n \log^2(n))$ time. *SIAM Journal on Computing*, 12(1):71–81, 1983.

[Sti93] J Stillwell. *Classical topology and combinatorial group theory.* Springer-Verlag, New York, second edition, 1993.

[TSN96] J. Takahashi, H. Suzuki, and T. Nishizeki. Shortest noncrossing paths in plane graphs. *Algorithmica*, 16:339–357, 1996.

[Tut01] W.T. Tutte. *Graph Theory.* Cambridge University Press, Cambridge, 2001.

[TV98] S. P. Tarasov and M. N. Vyalyi. Construction of contour trees in 3D in $O(n \log n)$ steps. In *Proceedings of the 14th Annual Symposium on Computational Geometry*, pages 68–75, Minneapolis, MN, 1998. ACM.

[VY90] G. Vegter and C. K. Yap. Computational complexity of combinatorial surfaces. In *Proceedings of the 6th Annual Symposium on Computational Geometry*, pages 102–111, Berkeley, CA, 1990. ACM.