Splitting (Complicated) Surfaces Is Hard[*]

Erin W. Chambers[†]   Éric Colin de Verdière[‡]   Jeff Erickson[§]

Francis Lazarus[¶]   Kim Whittlesey[‖]

Abstract

Let $\mathcal{M}$ be an orientable surface without boundary. A cycle on $\mathcal{M}$ is *splitting* if it has no self-intersections and it partitions $\mathcal{M}$ into two components, neither homeomorphic to a disk. In other words, splitting cycles are simple, separating, and non-contractible. We prove that finding the shortest splitting cycle on a combinatorial surface is NP-hard but fixed-parameter tractable with respect to the surface genus. Specifically, we describe an algorithm to compute the shortest splitting cycle in $g^{O(g)} n^2 \log n$ time.

Without knowing what futurism is like, Johansen achieved something very close to it when he spoke of the city; for instead of describing any definite structure or building, he dwells only on broad impressions of vast angles and stone surfaces - surfaces too great to belong to anything right or proper for this earth, and impious with horrible images and hieroglyphs. I mention his talk about angles because it suggests something Wilcox had told me of his awful dreams. He said that the geometry of the dream-place he saw was abnormal, non-Euclidean, and loathsomely redolent of spheres and dimensions apart from ours. Now an unlettered seaman felt the same thing whilst gazing at the terrible reality.

— H. P. Lovecraft, "The Call of Cthulhu" (1926)

---

1   Introduction

Optimization problems on surfaces in the fields of computational topology and topological graph theory have received much attention in the past few years. Such problems are usually set in the *combinatorial surface* model. A combinatorial surface is a graph $G(\mathcal{M})$ embedded on a surface $\mathcal{M}$ that cuts $\mathcal{M}$ into topological disks. Curves on this surface are required to be walks on $G(\mathcal{M})$, and edges of $G(\mathcal{M})$ have positive weights, allowing one to measure the length of a curve.

Most of these results have been obtained in the *combinatorial surface* model. Erickson and Har-Peled [8] prove that the computation of a shortest subgraph of $G(\mathcal{M})$ that cuts $\mathcal{M}$ into a disk, or shortest cut graph, is NP-hard. Colin de Verdière and Lazarus consider the problem of finding the shortest *simple* loop [5] or cycle [4] within a given homotopy class. A polynomial time algorithm for the generalization of this problem to non-simple curves was later obtained by Colin de Verdière and Erickson [3]. Erickson and Whittlesey [9] provide simple polynomial time algorithms to compute the shortest homology basis and the shortest fundamental system of loops of a surface. Cabello and Mohar [2] describe algorithms to compute the shortest non-separating and non-contractible cycle, improving earlier algorithms of Erickson and Har-Peled [8] when the genus of the input surface is small.

Many of these problems can be seen as the computation of a shortest cycle with some prescribed topological property, such as non-contractibility. When the set of cycles with the desired property satisfies the so-called *3-path condition*, a generic algorithm of Mohar and Thomassen finds a shortest such cycle in $O(n^3)$ time [14, Sect. 4.3]. For instance, the sets of non-separating and non-contractible cycles satisfy the 3-path condition, but the set of separating cycles on a surface does not.

In this paper, we study the following optimization problem: Given a orientable 2-manifold $\mathcal{M}$ with genus $g \geq 2$ without boundary, we find a shortest simple non-contractible cycle that separates $\mathcal{M}$. For simplicity, we will call a simple non-contractible separating cycle a *splitting* cycle. Removing a splitting cycle from any surface leaves two non-contractible components, each with one boundary cycle. Cutting along a maximal set of non-homotopic splitting cycles decomposes the surface into a collection of punctured tori.

After reviewing a few necessary concepts from topology and proving some preliminary results, we prove in Section 3 that it is NP-hard to compute a shortest splitting cycle. In Section 4, we then prove that a splitting cycle on a surface of genus g cuts any shortest path on the surface $O(g)$ times, and that this bound is tight in the worst case. This property leads to an algorithm to compute a shortest splitting cycle in $g^{O(g)}n^2 \log n$ time, which we describe in Section 5. Thus, the shortest splitting problem is fixed-parameter tractable with respect to the genus of the surface. This is the first result of this kind among the previously cited works. In particular, although Erickson and Har-Peled provide an polynomial-time algorithm to compute the minimum cut graph on any surface of constant genus, the order of the polynomial running time depends on the genus [8].

2   Preliminaries

2.1   Topological Background

We recall several notions from combinatorial and computational topology. See also Hatcher [11], Stillwell [15], or Zomorodian [16] for more details.

Curves on surfaces.   A *surface* (or 2-manifold with boundary) $\mathcal{M}$ is a topological Hausdorff space where each point has a neighborhood homeomorphic either to the plane or to the closed half-plane. The points with neighborhood homeomorphic to the closed half-plane comprise the *boundary* of $\mathcal{M}$. All the surfaces considered here are *compact*, *connected*, and *orientable*. Such a surface is homeomorphic to a sphere with g handles attached and b open disks removed, for some

unique non-negative integers g and b called respectively the *genus* and the *number of boundaries* of $\mathcal{M}$. In this paper, all surfaces are without boundary, unless specifically stated otherwise.

Let $\mathcal{M}$ be a surface. A *path* on $\mathcal{M}$ is a continuous map $p : [0,1] \to \mathcal{M}$. The *endpoints* of p are $p(0)$ and $p(1)$. A *loop* with *basepoint* x is a path with both endpoints equal to x. A *cycle* is a continuous map $\gamma : S^1 \to \mathcal{M}$, where S^1 denotes the unit circle. A path, loop, or cycle is *simple* if it is one-to-one, except, of course, for the endpoints of a loop. Two cycles are *disjoint* if they do not intersect. Two paths are *disjoint* if they intersect only at their endpoints; in particular, two disjoint loops intersect only at their common basepoint.

If p is a path, its *reversal* is the path $\bar{p}(t) = p(1 - t)$. The *concatenation* $p \cdot q$ of two paths p and q with $p(1) = q(0)$ is defined by setting $(p \cdot q)(t) = p(2t)$ for all $t \le 1/2$, and $(p \cdot q)(t) = q(2t - 1)$ for all $t \ge 1/2$.

Systems of loops and homotopy.   If L is a set of pairwise disjoint simple loops, $\mathcal{M} \setminus L$ denotes the surface with boundary obtained by *cutting* $\mathcal{M}$ along the loops in L. A *system of loops* on $\mathcal{M}$ is a set of pairwise disjoint simple loops L such that $\mathcal{M} \setminus L$ is a topological disk. Any system of loops contains exactly $2g$ loops. $\mathcal{M} \setminus L$ is a $4g$-gon where the loops appear in pairs, and is called the *polygonal schema* associated to L.

A *homotopy* between two paths p and q is a continuous map $h : [0,1] \times [0,1] \to \mathcal{M}$ such that $h(0, \cdot) = p$, $h(1, \cdot) = q$, $h(\cdot, 0) = p(0) = q(0)$, and $h(\cdot, 1) = p(1) = q(1)$. A (free) *homotopy* betweeen two cycles γ and δ is a continuous map $h : [0,1] \times S^1$ such that $h(0, \cdot) = \gamma$ and $h(1, \cdot) = \delta$. A loop or cycle is *contractible* if it is homotopic to a constant loop or cycle. A simple cycle γ is *separating* if $\mathcal{M} \setminus \gamma$ has two components. Every contractible simple cycle is separating (actually, it bounds a disk), but not all simple cycles are contractible.

Two cycles are *homologous* if one can be continuously deformed into the other via a deformation that may include splitting cycles at self-intersection points, merging intersecting pairs of cycles, or adding or deleting contractible cycles. For any two loops ℓ and ℓ' with the same basepoint, their concatenations $\ell \cdot \ell'$ and $\ell' \cdot \ell$ are homologous. A cycle or loop is *null-homologous* if it is homologous to a constant loop; a cycle is separating if and only if it is simple and null-homologous.

We say that a cycle *splits* a surface $\mathcal{M}$ if it is simple, non-contractible, and separating.

Combinatorial and cross-metric surfaces.   Like most earlier related results [2, 4, 5, 8, 9, 13], we state and prove our results in the *combinatorial surface* model. A combinatorial surface is an abstract surface $\mathcal{M}$ together with a weighted undirected graph $G(\mathcal{M})$, embedded on $\mathcal{M}$ so that each open face is a disk. In this model, the only allowed paths are walks in G; the length of a path is the sum of the weights of the edges traversed by the path, counted with multiplicity. The *complexity* of a combinatorial surface is the total number of vertices, edges, and faces of G.

It is often more convenient to work in an equivalent dual formulation of this model introduced by Colin de Verdière and Erickson [3]. A *cross-metric surface* is also an abstract surface $\mathcal{M}$ together with an undirected weighted graph $G^* = G^*(\mathcal{M})$, embedded so that every open face is a disk. However, now we consider only *regular* paths and cycles on $\mathcal{M}$, which intersect the edges of G^* only transversely and away from the vertices. The *length* of a regular curve p is defined to be the sum of the weights of the dual edges that p *crosses*, counted with multiplicity. See [3] for further discussion of these two models.

2.2   Preliminary Lemmas

A set of loops, all with the same basepoint, are *independent* if they are simple, non-contractible, pairwise non-homotopic, and pairwise disjoint. An independent set L of loops is *strongly independent* if no loop in L is homotopic to the reverse of another loop in L.

Lemma 2.1. *Let $\mathcal{M}$ be an orientable surface of genus $g \geq 1$. Any independent set of loops in $\mathcal{M}$ contains at most $12g - 6$ loops.*

Proof: Let L be a strongly independent set of loops with basepoint x; it suffices to show that $|L| \leq 6g - 3$. Each component of $\mathcal{M} \setminus L$ has at least one boundary cycle, and each boundary cycle contains at least one copy of the basepoint x. If any component $\mathcal{M}'$ is a disk, that disk must have at least three copies of the basepoint x on its boundary, because no loop in L is contractible or homotopic to another loop in L or its inverse. We will add loops to the set L, maintaining its strong independence, until we reach a maximal strongly independent set.

First, if some component $\mathcal{M}'$ has non-zero genus, let ℓ be a simple non-separating loop in $\mathcal{M}'$ based at (some copy of) x. $\mathcal{M}' \setminus \ell$ has smaller genus than $\mathcal{M}'$, and $L \cup \{\ell\}$ is still a strongly independent set of loops in $\mathcal{M}$.

Second, if some component $\mathcal{M}'$ has more than one boundary cycle, let ℓ' be a path in $\mathcal{M}'$ between any two copies of x on different boundary cycles. $\mathcal{M}' \setminus \ell'$ has the same genus but one fewer boundary than $\mathcal{M}'$. This path corresponds to a loop ℓ in $\mathcal{M}$, and $L \cup \{\ell\}$ is still strongly independent.

Finally, suppose every component $\mathcal{M}'$ of $\mathcal{M} \setminus L$ is homeomorphic to a disk. As we argued earlier, each component has at least three copies of x on its boundary. If four or more copies of x appear on the boundary of some component $\mathcal{M}'$, we can triangulate $\mathcal{M}'$ by adding more loops to L without violating strong independence.

Thus, if L is a *maximal* strongly independent set of loops, each component of $\mathcal{M} \setminus L$ is a triangle, that is, a disk with exactly three copies of x on its boundary. Let t be the number of triangles. Each triangle is incident to three loops, and each loop to two triangles, so $3t = 2|L|$, and Euler's formula implies that $1 - |L| + t = 2 - 2g$. We conclude that $|L| = 6g - 3$. $\qquad\square$

The intersection of any loop ℓ with a small neighborhood of the basepoint consists of an *initial subpath* ℓ^+ and a *terminal subpath* ℓ^-. Let L be a set of loops with common basepoint. The *incidence pattern* of L records the cyclic order of its initial and terminal subpaths around the basepoint.

Lemma 2.2. *Let a and b be two simple, non-contractible, disjoint, homotopic loops on an orientable surface $\mathcal{M}$. For one of the two possible orientations of $\mathcal{M}$, the incidence pattern of these loops is $a^+a^-b^-b^+$. Moreover, a and b bound an open disk whose intersection with a neighborhood of the basepoint consist of a wedge between a^+ and b^+ and a wedge between b^- and a^-.*
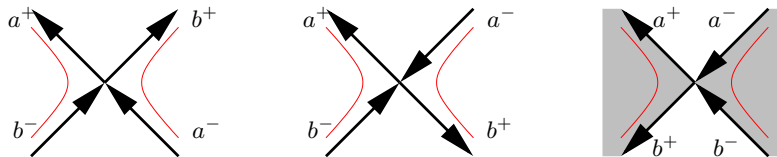

Figure 1. Lemma 2.2. Only the third incidence pattern is possible; a and b bound a disk containing the grey wedges.

Proof: If the incidence pattern were $a^+b^+a^-b^-$ or $a^+a^-b^+b^-$, there would be a simple cycle homotopic to the loop $a \cdot b \simeq a \cdot a$. (See Figure 1.) But this is impossible, because no loop homotopic to the square of a non-contractible simple loop is simple [7, Theorem 4.2]. So the incidence pattern must be $a^+a^-b^-b^+$.

Now, consider a contractible simple cycle γ that is a slightly translated copy of $a \cdot b^{-1}$ and intersects neither a nor b. In the neighborhood of the basepoint, γ is inside the wedge between a^+ and b^+ on one side, and inside the wedge between b^- and a^- on the other side. The disk

bounded by γ cannot contain the basepoint, because it would then contain the non-contractible loops a and b. Thus, in the neighborhood of the basepoint, this disk also lies inside the same two wedges. $\square$

3   NP-hardness

Theorem 3.1. *Finding the shortest splitting cycle on a combinatorial surface is NP-Hard.*

Proof: We describe a reduction from the following special case of the Euclidean traveling salesman problem [12]: *Given a set P of n points on the two-dimensional integer grid, are they connected by a tour whose length is exactly n?* Any such tour must lie entirely on the grid, and must consist of n axis-parallel unit-length segments, each joining a pair of points in P. Our reduction is similar to (and was inspired by) a proof by Eades and Rappaport [6] that computing the minimum-perimeter polygon separating a set of red points from a set of blue points in the plane is NP-hard; see also Arkin *et al.* [1].

     We describe a two-step reduction. Let P be a set of n points on the $n \times n$ integer grid in the plane. To begin the first reduction, we overlay n 4×4 square grids of width $\varepsilon < 1/4n$, one centered on each point in P, on top of the $n \times n$ integer grid. In each small grid, we color the square in the second row and second column *red* and the square in the third row and third column *blue*. We now easily observe that the following question is NP-complete: *Does the modified grid contain a cycle of length at most $n + 1/2$ that separates the red squares from the blue squares?* Any TSP tour of P of length n can be modified to produce a separating cycle of length at most $n + 1/2$ by locally modifying the tour within each small grid, as shown in Figure 2(a) and (b). Conversely, any separating cycle must pass through the center points of all n small grids, which implies that any separating cycle of length at most $n + 1/2$ must contain n grid edges that comprise a TSP tour of P.
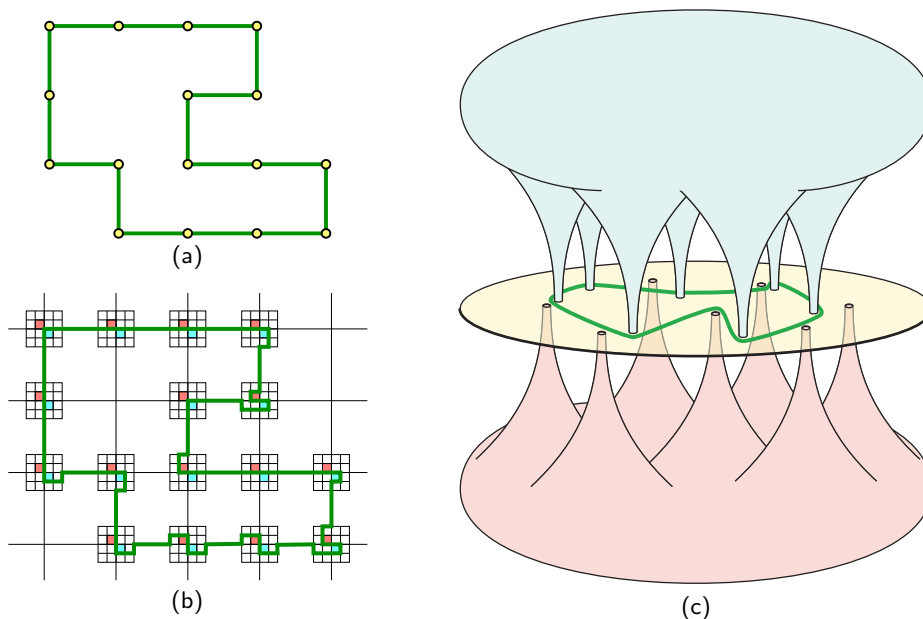

Figure 2. (a) A TSP tour of length n. (b) The corresponding red/blue separating cycle (not to scale). (c) Separating heaven from hell (not to scale); the central disk is a small portion of Earth.

     In the second reduction, we reduce the problem to finding a minimum-length splitting cycle. We isometrically embed the modified grid on a sphere, which we call *Earth*. We remove the red and blue

squares to create $2n$ punctures, which we attach to two new punctured spheres, called *heaven* and *hell*. We attach the n punctures in heaven to the n blue punctures on Earth; similarly, we attach the n punctures in hell to the n red punctures on Earth. We append edges of length n^3 to the resulting surface so that each face of the final embedded graph is a disk. The resulting combinatorial surface $\mathcal{M}(P)$ has genus $2n - 2$ and complexity $O(n^2)$, and it can clearly be constructed in polynomial time. See Figure 2(c).

The shortest cycle γ that splits $\mathcal{M}(P)$ must lie entirely on Earth, since the edges in heaven and hell are far too long. Moreover, γ must separate the blue punctures from the red punctures; otherwise, $\mathcal{M}(P) \setminus \gamma$ would be connected by a path through heaven or through hell. Thus, γ is precisely the shortest cycle that separates the red and blue squares in our intermediate problem. This cycle has length at most $n + 1/2$ if and only if the original points P have a tour of length n. $\quad\square$

4 Structural Properties

For any two points x and y on a cycle α, we let $\alpha[x, y]$ denote the path from x to y along α, taking into account the orientation of α. For a path or a dual edge α, the same notation is used for the unique simple path between x and y on α.

4.1 Multiplicity Bound

Lemma 4.1. *Any shortest splitting cycle on a cross-metric surface $\mathcal{M}$ crosses each edge e^* of $G^*(\mathcal{M})$ at most once in each direction.*

Proof: Assume for the purpose of contradiction that some shortest splitting cycle γ crosses some dual edge e^* twice in the same direction, say left to right. Let x and z be consecutive left-to-right intersection points along e^*; that is, γ does not cross $e^*[x, z]$ from left to right. The path $e^*[x, z]$ is on one side of γ near x, but on the other side of γ near z. Because γ separates the surface, γ must cross e^* from right to left at some point y between x and z.

The cycles $\gamma[x, y] \cdot e^*[y, x]$ and $\gamma[y, z] \cdot e^*[z, y]$ are non-contractible; otherwise, we could shorten γ by removing two crossings with e^* without changing its homotopy class.
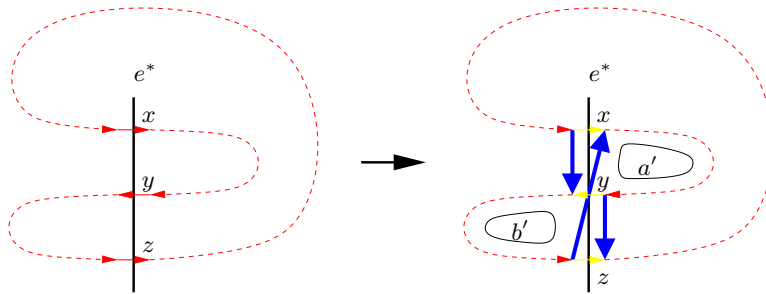

Figure 3. Lemma 4.1. If γ crosses e^* twice in the same direction, we can remove both crossings.

Define a new cycle $\gamma' = \gamma[x, y] \cdot e^*[y, z] \cdot \gamma[z, x] \cdot e^*[x, y] \cdot \gamma[y, z] \cdot e^*[z, x]$, as shown in Figure 3. The new cycle γ' is simple, because x, y, z are consecutive along e^*. The cycle γ' is in the same (integer) homology class as γ and is therefore null-homologous. By translating the non-contractible cycles $\gamma[x, y] \cdot e^*[y, x]$ and $\gamma[y, z] \cdot e^*[z, y]$ away from γ', we obtain two non-contractible cycles a' and b', one in each component of $\mathcal{M} \setminus \gamma'$; so γ' is non-contractible. Finally, γ' crosses e^* two times fewer than γ and crosses every other edge in $G^*(\mathcal{M})$ the same number of times as γ. We conclude that γ' is a splitting cycle that is shorter than γ, which is impossible. $\quad\square$

4.2 $O(g)$ Crossings with Any Shortest Path

Proposition 4.2. *Let P be a set of pairwise interior-disjoint shortest paths on a cross-metric surface $\mathcal{M}$. Some shortest splitting cycle crosses each path in P at most $O(g)$ times.*

Proof: Let γ be a shortest splitting cycle with the minimum number of crossings with paths in P. We can assume that γ does not pass through the endpoints of any path in P, since we are on a cross-metric surface and could simply perturb γ slightly. Consider any path p in P that intersects γ.

The intersection points $\gamma \cap p$ partition γ into *arcs*. These arcs may intersect other paths in P. Let $\mathcal{M}/p$ be the quotient surface obtained by contracting p to a point p/p. Each arc corresponds to a loop in $\mathcal{M}/p$ with basepoint p/p. We say that two such arcs are *homotopic rel p* or *relatively homotopic* if the corresponding loops in $\mathcal{M}/p$ are homotopic.

For any two consecutive intersection points x and y along γ, the arc $\gamma[x,y]$ cannot be homotopic to $p[x,y]$, since otherwise we can obtain a no longer splitting cycle $\gamma[y,x] \cdot p[x,y]$ that has fewer crossings with the paths in P. It follows that none of the arcs of γ are contractible rel p. Since the arcs are disjoint except at their common endpoints, Lemma 2.1 implies that there are at most $12g - 6$ relative homotopy classes of arcs.

We can partition the arcs into four types—LL, RR, LR, and RL—according to whether the arcs start on the left or right side of p, and whether they end on the left or right side of p. To complete the proof, we argue that there is at most one arc of each type in each relative homotopy class. We explicitly consider only types LL and RL; the other two cases follow from symmetric arguments.

Suppose for purposes of contradiction that there are two LL-arcs $u = \gamma[a,z]$ and $w = \gamma[c,x]$ that are homotopic rel p. By Lemma 2.2, without loss of generality, the intersection points appear along e^* in the order a, c, x, z, and the cycle $u \cdot p[z,x] \cdot \bar{w} \cdot p[c,a]$ bounds a disk. Without loss of generality, we assume that no other arc homotopic rel p intersects this disk. Since γ is separating, ~~there must be~~ exactly one arc $v = \overline{\gamma[g,b]}$ between ~~a~~ and w that is relatively homotopic to $\bar{u}$ and $\bar{w}$.
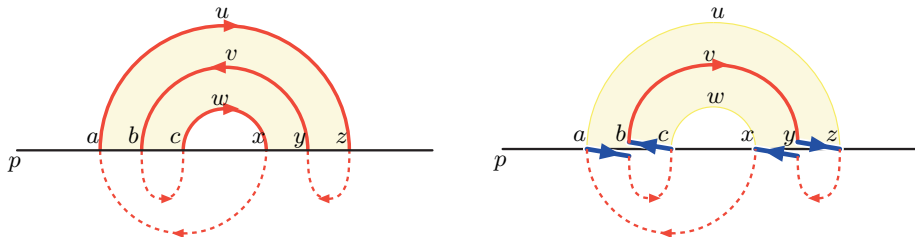

Figure 4. The exchange argument for LL arcs.

Without loss of generality, suppose the path $\gamma[x,a]$ does not contain any of the arcs u, v, or w. Consider the cycle

$$\gamma' = p[a,b] \cdot \gamma[b,c] \cdot p[c,b] \cdot \bar{\gamma}[b,y] \cdot p[y,z] \cdot \gamma[z,y] \cdot p[y,x] \cdot \gamma[x,a]$$

obtained by removing u and w from γ, reversing v, and connecting the remaining pieces of γ with subpaths of p; see Figure 4. This cycle crosses p fewer times than γ, and crosses any other path in p no more than γ. An argument identical to the proof of Lemma 4.1 implies that γ' is simple, null-homologous, and non-contractible. Because p is a shortest path, u cannot be shorter than $p[a,z]$, which implies that γ' is no longer than γ, contradicting the fact that γ is a shortest splitting cycle with the minimal number of crossings with paths in P. We conclude that any two LL-arcs must be in different relative homotopy classes.

Now suppose for purposes of contradiction that there are two RL-arcs $u = \gamma[a, x]$ and $w = \gamma[c, z]$ that are relatively homotopic. By Lemma 2.2, the cycle $u \cdot p[x, z] \cdot \bar{w} \cdot p[c, a]$ bounds a disk D. Without loss of generality no arc relatively homotopic to u and w lies inside this disk. Up to symmetry, we can assume that a precedes c and x precedes z along edge e^*. Point x cannot lie between a and c, because then arc u would end on the boundary of D, and γ could not exit D without creating a self-intersection, an arc that is contractible rel p, or an arc in D that is relatively homotopic to u and w, none of which are possible. Thus, the four intersection points must appear in the order a, c, x, z, possibly with $c = x$. See Figure 5.
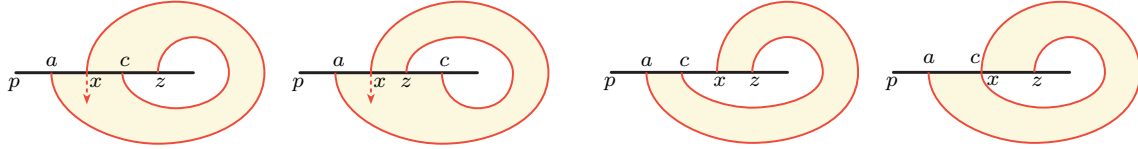

Figure 5. Two impossible incidence orders and two possible incidence orders for two RL arcs; disk D is shaded.

As in the previous case, because γ is a splitting cycle, there must be exactly one arc $v = \gamma[y, b]$ between u and w that is homotopic to $\bar{u}$ and $\bar{w}$. If $c = x$, there is only one way to connect these three arcs to form the cycle γ; if $c \neq x$, there are two possibilities. See Figure 6. In all three cases, by deleting arcs u and w, reversing arc v, and connecting the remaining pieces of γ with subpaths of p, we create a splitting cycle γ' that is no longer than γ and crosses fewer times the paths in P than γ, which is impossible. We omit the tedious details. We conclude that any two RL-arcs must be in different relative homotopy classes. □
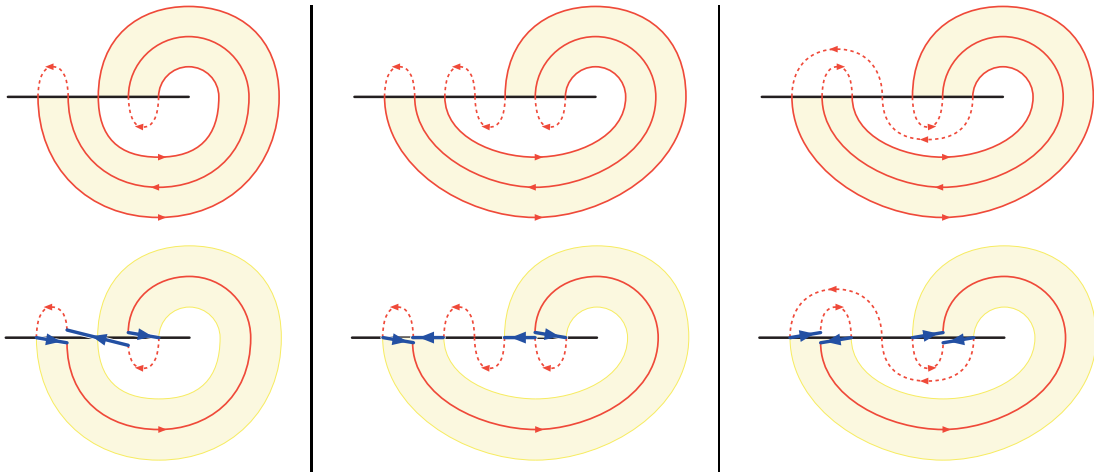

Figure 6. Three exchange arguments for RL arcs

4.3   Shortest Splitting Cycles Can Be Complicated

Earlier algorithms for computing shortest cycles with some specific topological property (non-contractible, non-separating, or essential) rely on the fact that the desired cycle is the concatenation of two equal-length shortest paths [14, 8].* Cabello and Mohar [2] exploit a slightly different property to compute shortest nontrivial cycles more quickly on surfaces of constant genus: the desired shortest cycle crosses any shortest path a constant number of times. As we prove next,

---

*In general, this characterization requires shortest paths that terminate in the interior of edges, but if we refine edges appropriately, the shortest cycle will indeed be the concatenation of two shortest vertex-to-vertex paths.

neither of these properties holds for the shortest splitting cycle; in particular, the upper bound of Proposition 4.2 is tight up to constant factors.

Theorem 4.3. *For any $g \geq 2$, there is a combinatorial surface $\mathcal{M}_g$ of genus g whose unique shortest splitting cycle (up to orientation) crosses a shortest path $\Omega(g)$ times and (therefore) cannot be decomposed into fewer than $\Omega(g)$ shortest paths and edges.*

Proof: We consider only the case where g is even. We construct a combinatorial surface $\mathcal{M}_g$ of genus g as follows; see Figure 7. The base surface $\mathcal{M}_0$ is a sphere whose geometry approximates an hourglass. Let γ be the central cycle that partitions the two lobes of the hourglass. To construct $\mathcal{M}_g$, we attach g handles to this hourglass, each joining a small circle c_i on the neck of the hourglass, just to one side of γ, to a large circle C_i far away on the opposite lobe. The small circles c_i are arranged symmetrically around the neck of the hourglass, alternating between the two sides of γ; the large circles C_i are also partitioned evenly between the two lobes of the hourglass.

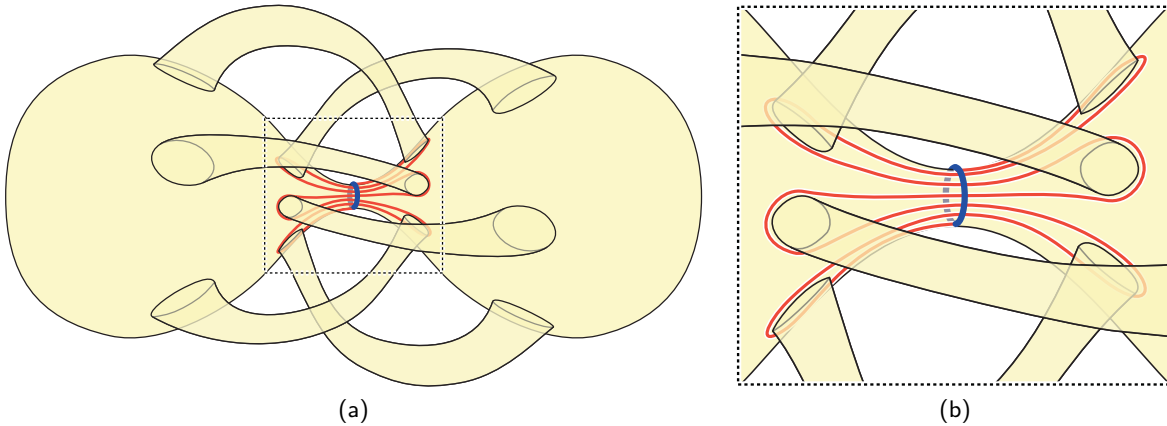(a)   (b)

Figure 7. Ph'nglui mglw'nafh Cthulhu R'lyeh wgah'nagl fhtagn! (a) A surface whose shortest splitting cycle cuts a shortest path $\Omega(g)$ times. (b) A closeup of the undulating shortest splitting cycle.

Let μ be a cycle that undulates around the small circles in order, crossing γ a total of g times, as shown in Figure 7(b). We easily verify that μ splits $\mathcal{M}_g$ into two surfaces of genus g. Let $x_1, x_2, \ldots x_g = x_0$ denote the g intersection points of μ and γ. For each i, let μ_i and γ_i respectively denote the subpaths of μ and γ between x_{i-1} and x_i. Let F be the union of these $2g$ paths.

To obtain a combinatorial structure on $\mathcal{M}_g$, we embed a weighted graph G that contains F as a subgraph, such that every face of the embedding is a topological disk. We assign each edge μ_i length $2g$, each path γ_i length 1, and every other edge in G length at least $4g^2$. With the assigned edge weights, any shortest splitting cycle must be a circuit in F.

Let α be a splitting cycle that is a circuit in F. By Lemma 4.1, α traverses each path γ_i and μ_i at most once in each direction. For any i, there is a path in $\mathcal{M}_g \setminus F$ from one side of γ_i to the other, so α must traverse each edge γ_i either twice (in opposite directions) or not at all.

Consider any simple cycle β in a tubular neighborhood of F that traverses every edge in F either once in each direction or not at all. This cycle must be null-homologous, and therefore separating. Since β lies inside a small tubular neighborhood of F, which has genus zero, some component of $\mathcal{M} \setminus \beta$ has genus zero. Moreover, β is the only boundary of this component. We conclude that β is contractible.

The splitting cycle α is not contractible, so it must traverse some edge μ_i exactly once. But α must traverse the edges adjacent to any vertex x_i an even number of times, which implies that α traverses *every* edge μ_i exactly once. Thus, every splitting cycle in F is at least as long as μ, which implies that μ is the unique shortest splitting cycle in $\mathcal{M}_g$.

The cycle μ does not contain a single (even approximate!) vertex-to-vertex shortest path. Even if we allow shortest paths between points in the interior of edges, each such path contains at most one vertex x_i. The cycle γ can clearly be partitioned into two shortest paths of length $g/2$; μ crosses each of these paths at least $g/2 - 2$ times. $\square$

5   Algorithm

In this section, we prove the following theorem.

Theorem 5.1. *Let $\mathcal{M}$ be an orientable cross-metric surface without boundary; let g be its genus and n be its complexity. We can compute a shortest splitting cycle in $\mathcal{M}$ in $g^{O(g)}n^2 \log n$ time.*

The algorithm proceeds in several stages, described in detail in the following subsections. First, we compute the shortest system of loops from some arbitrary basepoint, using the greedy algorithm of Erickson and Whittlesey [9]. Next, we cut $\mathcal{M}$ along this system of loops and then enumerate all possible homotopy types for the splitting cycle on the resulting polygonal schema. We discard any homotopy type that does not yield a valid splitting cycle. Finally, for each remaining homotopy class, we compute a representative cycle γ on the input surface $\mathcal{M}$, and then compute the shortest cycle homotopic to γ, using the recent algorithm of Colin de Verdière and Erickson [3]. Out of all cycles constructed this way, we return the shortest one.

5.1   Greedy Loops

Let v be any point of $\mathcal{M}$ in the interior of a face of $G^*(\mathcal{M})$. Let $\alpha_1, \ldots, \alpha_{2g}$ be the shortest system of loops of $\mathcal{M}$ with basepoint v; this system of loops can be computed in $O(n \log n + gn)$ time using a greedy algorithm of Erickson and Whittlesey [9].

The key property of this system of loops is that each loop α_i is composed of two shortest paths in the primal graph $G(\mathcal{M})$; however, in general these two paths meet at a point m_i in the interior of some edge e_i. To simplify our algorithm, we split e_i into two edges at m_i—or equivalently, in the dual graph, we replace the dual edge e_i^* with two parallel edges—partitioning the length appropriately, so that α_i consists of two vertex-to-vertex shortest paths β_i and β_i' in $G(\mathcal{M})$.

5.2   Enumeration of Homotopy Classes via Labeled Triangulations

Lemma 4.2 tells us that some shortest splitting cycle γ that crosses each path β_i or β_i' at most $O(g)$ times, and thus crosses each loop α_i at most $O(g)$ times. Our algorithm therefore enumerates (possibly redundantly) a superset of all homotopy classes of cycles that cross each α_i at most $O(g)$ times.

Cut $\mathcal{M}$ along the loops α_i to obtain a polygonal schema. This operation also cuts the unknown cycle γ into many segments that go across the schema. Since γ is simple, no two of these segments cross. We can assume that γ does not pass through the basepoint v; since v does not lie on $G^*(\mathcal{M})$, we can slightly perturb γ without changing its length (in the cross-metric surface) or its homotopy class.

The segments of γ can be grouped into subsets according to which pair of edges they meet on the polygonal schema (Figure 8). We dualize the polygonal schema, replacing each edge with a vertex and connecting vertices that correspond to consecutive edges; now each subset of segments corresponds to an edge between two vertices of the dual $4g$-gon. Since no two segments cross, these edges cannot cross; in particular, all the edges belong to some triangulation of the dual polygon.

Thus the candidate homotopy classes of a shortest splitting cycle are described by *labeled triangulations*, which consists of a triangulation of the dual polygon, in which every edge is labeled
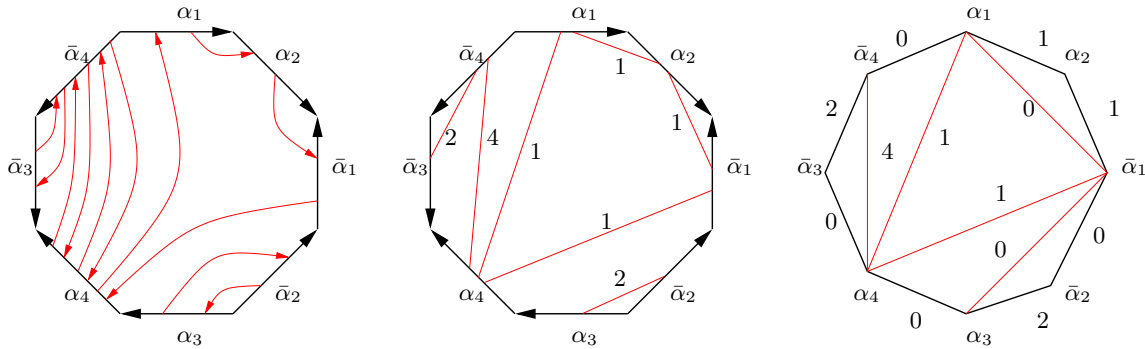
Figure 8. From a cycle to a labeled triangulation. Left: A splitting cycle. Middle: The corresponding subsets of segments; each label indicates the number of segments contained in a subset. Right: A corresponding labeled triangulation of the dual polygon.

with an integer between 0 and $O(g)$. Intuitively, the label of an edge in the triangulation represents the number of times that the cycle runs along that edge. There are $C_{4g-2} = O(4^{4g})$ possible triangulations, where C_n is the nth Calatan number, which we can enumerate in $O(g)$ time each. (This is essentially identical to the enumeration of binary trees.) There are $g^{O(g)}$ ways to label each triangulation, which we can enumerate in constant amortized time per labeling.

We thus obtain a total of $g^{O(g)}$ potential homotopy classes for γ. Most of the labeled triangulations do not correspond to a splitting cycle, or to any cycle for that matter. We now explain how to throw away a number of these possibilities.

5.3   Discarding Irrelevant Labeled Triangulations

Given a candidate labeled triangulation T, we want to test whether this corresponds to a splitting cycle. We must check that this representation (1) corresponds to a set of cycles, (2) is actually a single cycle, (3) is separating and (4) is non-contractible. We describe how to perform these tests in $O(g^2)$ time.

To check that T corresponds to a set of cycles, it suffices to check that two edges in the polygonal schema that correspond to the same α_i are crossed the same number of times. If this property is not satisfied, we simply discard T.

For the three following steps, we build a combinatorial surface corresponding to the arrangement, on $\mathcal{M}$, of the greedy loops and of the candidate cycle(s) defined by T. Imagine cutting the polygonal schema along the subsets of edges given by the triangulation. (Since we have multiple edges actually running along one edge of the triangulation, we view these as giving thin strips between triangular pieces.) Then identify corresponding subpaths of the polygonal schema. The complexity of the resulting surface is $O(g^2)$. Said differently, this surface is exactly the surface $\mathcal{M}$ containing the arrangement of the greedy system of loops and of the candidate cycle, but forgetting about the $O(n)$ internal complexity of the surface.

We can now check that T defines a single cycle by walking along the graph, starting from an arbitrary segment, marking visited segments until we return to the initial segment, and then by checking that all segments have been marked.

To test that a cycle is separating but non-contractible, we use a simplification of an algorithm of Erickson and Har-Peled [8]. To test separation, perform a depth- or breadth-first search on the faces of the combinatorial surface, starting from any initial face, but forbidding crossings with any segment of the cycle. The cycle is separating if and only if the search halts before visiting every face of the surface. We also compute the Euler characteristic of the reachable portion of the surface

during the search, by counting vertices, edges, and faces. The cycle is non-contractible if and only if this Euler characteristic is neither 1 (a disk) nor $1 - 2g$ (the complement of a disk).

5.4   Shortest Cycle in Each Homotopy Class

In the last step, for each non-discarded labeled triangulation, we build a cycle in the corresponding homotopy class in the original cross-metric surface $\mathcal{M}$ and shorten it as much as possible in its homotopy class. We then retain a cycle with minimal length. Recall that, by the considerations above, there remain $g^{O(g)}$ labeled triangulations.

Let T be a labeled triangulation. We first create some paths in the polygonal schema that correspond to the interior edges of the triangulation. These paths are computed by running along a part of the polygonal schema. Specifically, we maintain a disk D inside which we need to create connections; initially, D is the polygonal schema. The dual of this disk is a tree. We iteratively remove a leaf of the tree from the disk and add a path corresponding to the edge of the triangulation that the leaf's edge crossed. The path is created by running along the boundary of the current disk. In particular, its complexity is $O(gn)$.

Now, we use the labels of T to create as many segments on each edge of the triangulation, and we connect these pieces in the obvious way. Since we have $O(g^2)$ paths, each consisting of $O(gn)$ edges, we construct a cycle γ with complexity $O(g^3 n)$ in $O(g^3 n)$ time.

We then use the algorithm of Colin de Verdière and Erickson [3] to find the shortest cycle γ' freely homotopic to γ; this algorithm runs in time $O(n^2 \log n + gnk \log(gnk))$, where k is the complexity of the input path. The input cycle γ has complexity $O(g^3 n)$, and trivially $g = O(n)$, so the running time simplifies to $O(g^4 n^2 \log n)$.

Finally, the output cycle γ' may contain self-intersections. However, because γ' is homotopic to a simple cycle, a theorem of Hass and Scott [10] (see also [4]) implies that if γ' is self-intersecting, some pair of self-intersections bounds a disk. We can discover any such bigon by depth-first search and then remove it by homotoping the cycle. This modification does not change the length of the cycle. The complexity of γ' is at most $2n$ by Lemma 4.1, so discovering and removing a bigon takes $O(n)$ time. Thus, in $O(n^2)$ time, we can remove all self-intersections from γ'.

Since our enumeration includes the homotopy class of the shortest splitting cycle, our algorithm will eventually output the shortest splitting cycle. The total time spent is $g^{O(g)} n^2 \log n$. This concludes the proof of Theorem 5.1.

6   Conclusions

The results of this paper suggest several open problems. Most notably, can we approximate the shortest splitting cycle, or is that also NP-hard? We could also consider the following high-level approach. Compute shortest simple cycles in each non-trivial homotopy class in order of increasing length, stopping either when we find a separating cycle, or when we find two cycles α and β that intersect an odd number of times. If we find a separating cycle, it is of course the shortest splitting cycle. If we find two cycles with odd intersection number, an exchange argument implies that the intersection number is 1. For some orientation of α and β, the cycle $\alpha \cdot \beta \cdot \bar{\alpha} \cdot \bar{\beta}$ is a splitting cycle whose length is at most four times the length of the shortest splitting cycle. Can this algorithm be implemented efficiently? How quickly can we enumerate the k shortest homotopy classes of (simple) cycles? Techniques of Eppstein for enumerating k shortest paths may be useful here.

If we iterately cut along splitting cycles, we optain a decomposition of the surface into punctured tori. Can we obtain the shortest such decomposition by repeatedly cutting the surface along its *shortest* splitting cycle? Is computing the shortest torus decomposition NP-hard? Similarly, a *pants decomposition* is a set of disjoint simple cycles decomposing a surface into pairs of pants, or spheres

with three boundary components [4]. Is it NP-hard to compute the shortest pants decomposition? Are either of these problems fixed-parameter tractable?

Finally, Erickson and Har-Peled [8] prove that finding the minimum cut graph is NP-hard, by a reduction from the fixed-parameter tractable rectilinear Steiner tree problem. Is computing the shortest cut graph fixed-parameter tractable? The most serious bottleneck here seems to be computing the shortest cut graph in a given homotopy class, where the homotopy is allowed to move the vertices.

References

[1] E. M. Arkin, S. Khuller, and J. S. B. Mitchell. Geometric knapsack problems. *Algorithmica* 10:399–427, 1993.

[2] S. Cabello and B. Mohar. Finding shortest non-separating and non-contractible cycles for topologically embedded graphs. *Proc. 13th Annu. European Sympos. Algorithms*, 131–142, 2005. Lecture Notes in Computer Science 3669, Springer-Verlag.

[3] É. Colin de Verdière and J. Erickson. Tightening non-simple paths and cycles on surfaces. *Proc. 17th Annu. ACM-SIAM Sympos. Discrete Algorithms*, p. to appear, 2006.

[4] É. Colin de Verdière and F. Lazarus. Optimal pants decompositions and shortest homotopic cycles on an orientable surface. *Proc. 11th Sympos. Graph Drawing*, 478–490, 2003. Lecture Notes Comput. Sci. 2912, Springer-Verlag.

[5] É. Colin de Verdière and F. Lazarus. Optimal systems of loops on an orientable surface. *Discrete Comput. Geom.* 33(3):507–534, 2005.

[6] P. Eades and D. Rappaport. The complexity of computing minimum separating polygons. *Pattern Recogn. Lett.* 14(9):715–718, 1993.

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

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

[9] J. Erickson and K. Whittlesey. Greedy optimal homotopy and homology generators. *Proc. 16th Annu. ACM-SIAM Sympos. Discrete Algorithms*, 1038–1046, 2005.

[10] J. Hass and P. Scott. Intersections of curves on surfaces. *Israel J. Math.* 51:90–120, 1985.

[11] A. Hatcher. *Algebraic topology.* Cambridge University Press, 2002. ⟨http://www.math.cornell.edu/~hatcher/⟩.

[12] D. S. Johnson and C. H. Papadimitriou. Computational complexity and the traveling salesman problem. *The Traveling Salesman Problem*, 68–74, 1985. John Wiley & Sons.

[13] F. Lazarus, M. Pocchiola, G. Vegter, and A. Verroust. Computing a canonical polygonal schema of an orientable triangulated surface. *Proc. 17th Annu. ACM Sympos. Comput. Geom.*, 80–89, 2001.

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

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

[16] A. Zomorodian. *Topology for Computing.* Cambridge University Press, 2005.