J. DIFFERENTIAL GEOMETRY
0 (2020) 1-19

CONSTRUCTING MONOTONE HOMOTOPIES AND SWEEPOUTS

Erin Wolf Chambers, Gregory R. Chambers, Arnaud de Mesmay, Tim Ophelders & Regina Rotman

Abstract

This article investigates when homotopies can be converted to monotone homotopies without increasing the lengths of curves. A monotone homotopy is one which consists of curves which are simple or constant, and in which curves are pairwise disjoint. We show that, if the boundary of a Riemannian disc can be contracted through curves of length less than L, then it can also be contracted monotonically through curves of length less than L. This proves a conjecture of Chambers and Rotman. Additionally, any sweepout of a Riemannian 2-sphere through curves of length less than L can be replaced with a monotone sweepout through curves of length less than L. Applications of these results are also discussed.

1. Introduction

The primary objects of study in this article are *monotone homotopies*, which we define below. Throughout the article, we consider closed curves on Riemannian surfaces. If α is a simple closed contractible curve, then $D(\alpha)$ denotes the closed disc that α bounds. If the underlying surface has at least one boundary component, then this disc is unique. If it is an oriented sphere, then the orientation of the sphere and the orientation of α determines $D(\alpha)$; it is the unique disc for which, given the orientation of the sphere, the induced orientation of the boundary agrees with that of α. If α and β are two simple closed contractible curves with $D(\beta) \subset D(\alpha)$, then let $A(\alpha,\beta) = A(\beta,\alpha)$ denote the annulus between α and β, that is, $D(\alpha)$ with the interior of $D(\beta)$ removed. If β is a constant curve, then we extend the definition of $A(\alpha,\beta)$ to denote $D(\alpha)$.

Definition 1.1. Let (M,g) be a Riemannian annulus with boundaries γ_0 and γ_1, and let $H : \mathbb{S}^1 \times [0,1] \to M$ be a homotopy between γ_0 and γ_1, that is, a smooth map such that $H(t,0) = \gamma_0$ and $H(t,1) = \gamma_1$. We will say that H is *monotone* if every intermediate curve $\gamma_\tau := H(t,\tau)$ is a simple closed curve parameterized by t, and if the closed 2-annuli

Received August 26, 2017.

$A(\gamma_\tau, \gamma_1) \subseteq M$ satisfy the inclusion $A(\gamma_{\tau_2}, \gamma_1) \subset A(\gamma_{\tau_1}, \gamma_1)$ for every $\tau_1 < \tau_2$. In this definition, γ_0 and γ_1 can be constant curves or simple closed curves.

A *monotone contraction* of a Riemannian 2-disc is a monotone homotopy from its boundary to a constant curve. We say that such a monotone homotopy is *outward* if $D(\gamma_0) \subset D(\gamma_1)$, and is called *inward* if $D(\gamma_1) \subset D(\gamma_0)$.

We prove the following two theorems, the first of which was a conjecture by Chambers and Rotman [**9**, Conjecture 0.2].

Theorem 1.2. *Suppose that (D, g) is a Riemannian disc, and suppose that there is a contraction of ∂D through curves of length less than L. Then there is a monotone contraction of ∂D through curves of length less than L.*

The techniques involved in the proof of this theorem also apply[1] in the setting of a Riemannian annulus and a homotopy between its two boundaries through curves of length less than L, yielding a monotone homotopy through curves of length less than L.

The second theorem concerns a similar monotonicity result for sweepouts of 2-spheres. A sweepout of a Riemannian 2-sphere is a map $f : S^1 \times S^1 \to S^2$ of degree 1. We can regard a sweepout as a 1-parameter family of connected closed curves $f(t, \cdot)$ parametrized by $t \in S^1$. These curves might have self-intersections as well as pairwise intersections.

Theorem 1.3. *Suppose that (S^2, g) is a Riemannian 2-sphere, and suppose that f is a sweepout of it composed of curves of length less than L. Then there exists a diffeomorphism from the round sphere $(S^2, round) = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ to (S^2, g) such that the length of the image of each parallel $\{(x, y, z) : z = constant\} \cap (S^2, round)$ is less than L.*

The proof of this result holds also if we assume only that there exists such a map of odd degree (which is not necessarily equal to 1).

Background and related work. These theorems have numerous applications to metric geometry, and to applied topology. In terms of metric geometry, Theorem 1.2 improves known estimates of the lengths of the shortest geodesics between pairs of points on Riemannian 2-spheres from [**11**] and [**13**]. In particular, the two authors prove that there are at least k geodesics joining any two points on a Riemannian 2-sphere of length at most $22kd$, where d is the diameter of the 2-sphere (if the two points agree, then this improves to $20kd$). These results improve these bounds to $16kd$ and $14kd$ respectfully, and also greatly decrease the complexity of the proofs in [**11**] and [**13**].

[1] The proof is even simpler in that case, since case b of Proposition 2.8 never occurs.

These results also allow the results from [**12**] to be generalized to the free loop space of a Riemannian 2-sphere; a map from $S^m \to \Lambda M$, where M is a Riemannian 2-sphere and ΛM is the set of closed curves in M, can be homotoped to a map $\tilde{f} : S^m \to \Lambda M$ consisting of curves of lengths bounded by $\varrho(m, k, d)$, with ϱ being an explicit function, d being the diameter of M, and k being the number of distinct non-trivial periodic geodesics of M with length at most $2d$. In [**12**], Nabutovsky and Rotman proved the analogous statement for maps into the space of simple closed curves based at a fixed point; our results allow this restriction to be removed. For more details, we refer to Chambers and Rotman [**9**, Section 0.1]. Of special note is that Theorem 1.2 directly implies that, if the boundary of a Riemannian disc is contractible through curves of length less than L, then for any point q on the boundary of the disc, the boundary is contractible to the point q through loops based at q of length less than $L + 2d$. Here, d is the diameter of the Riemannian 2-disc.

The sweepouts described in Theorem 1.3 appear in minimal surface and min-max literature. In [**8**], Chambers and Liokumovich show that if there is a sweepout of a Riemannian 2-sphere through curves of length less than L, then there is a sweepout of the same Riemannian 2-sphere through simple closed curves and constant curves of lengths less than L. They then use this result to answer a question of Freedman about the minmax levels with respect to different classes of sweepouts. Our theorem is an improvement on this result, proving that such a sweepout can be simplified to not only consist of curves which do not have self-intersections (other than constant curves), but to consist of such curves which are (mostly) pairwise disjoint as well.

From the computational topology literature, much recent work has focused on computing a "best" homotopy between two curves as a means of measuring similarity of the curves or determining optimal morphs between them [**4**, **6**, **10**]. The main goal in this setting is to determine the computational complexity of such a problem in the most common settings, generally where the two curves are in the plane (possibly with obstacles) or on a meshed surface, as typically produced by surface reconstruction algorithms.

The type of optimality we study in this work has been investigated in a combinatorial setting, where it was called the "height" of the homotopy [**2**, **3**, **10**], and in the graph theoretic setting, where it was called a "b-northward migration" [**1**]. However, the exact complexity of this problem remains open, and both papers include a conjecture that the best such morphings will proceed monotonically. The monotonicity result we present in this paper is a key ingredient in showing that this problem lies in the complexity class $\mathcal{NP}$ [**5**].

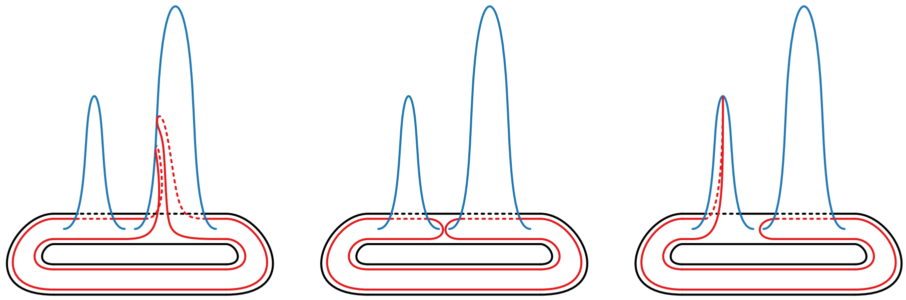

Figure 1. A counter-example to Conjecture 0.3 of Chambers and Rotman [9].

Finally, Chambers and Rotman formulated another conjecture [9, Conjecture 0.3] on monotonicity, where the initial curve is not the boundary of the disc. We say that a monotone contraction covers a simple closed curve γ if γ is contained in the disc which is the image of that contraction. They conjectured that if M is a Riemannian surface and γ a simple closed curve contractible through curves of length less than L, then there is a monotone contraction covering γ through curves of length less than L. We observe that this conjecture is false by exhibiting the counter-example in Figure 1.

In that example, the underlying surface is an annulus, and the metric is the Euclidean one, except for two mountains, one taller than the other one. The initial closed curve γ lies as shown in the first picture, half-way up the tall mountain from both sides. An optimal contraction is pictured in the following two pictures; it first climbs down the tall mountain *for both sides of the curve* in order to reduce the length, before climbing over the smaller one. On the other hand, any monotone contraction covering γ must start at a closed curve α that also lies half-way up the tall mountain from both sides. Then, by monotonicity, only one side of the curve can climb down the tall mountain. Therefore, the maximum length of the curves in such a monotone contraction will need to be strictly larger than for a non-monotone one.

Acknowledgments. This work was partially supported by NSF grants CCF-1614562 and CCF-1054779, NWO project no. 639.023.208, ANR project ANR-16-CE40-0009-01 (GATO), NSERC Discovery Grant RGPIN 217655-13, and NSERC Postdoctoral Fellowship PDF-487617-2016.

2. Preliminaries

We begin by recalling several definitions. A *Riemannian disc* is a closed smooth 2-dimensional Riemannian manifold with boundary that is diffeomorphic to a closed unit disc $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ with a smooth Riemannian metric; throughout this article, we will call this simply a "disc". A *Riemannian annulus* is a 2-dimensional smooth Riemannian manifold with boundary that is diffeomorphic to an annulus $\{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\}$ endowed with a smooth Riemannian metric; throughout this article, we will refer to such a manifold with boundary simply as an "annulus." A closed curve in a smooth manifold with boundary M is a smooth map from S^1 to M; a simple closed curve is a closed curve which is injective. An arc in a smooth manifold with boundary M is a smooth map from $[0, 1]$ to M; an arc of a closed curve in M is simply the restriction of the smooth map from S^1 to M to a closed subinterval of S^1. A homotopy of curves in a smooth manifold with boundary M is a smooth map $H : [0, 1] \times S^1 \to M$. We remark that, at several points in this article, we will form a new curve taking a curve and replacing an arc of that curve with a new arc. This will create at most two points which are not smooth, however, the resulting curve can be replaced by a smooth curve with length increased an arbitrarily small amount, and such that the new curve agrees with the old curve outside of balls of arbitrarily small radii centered at the two singular points. If the original curve is simple, then the new, smooth, curve is also simple. In this article, we implicitly assume that this smoothing procedure is executed whenever we execute such a cut-and-paste operation; we don't explicitly mention it to simplify the exposition.

Throughout the article, a closed curve γ in a Riemannian annulus A is called a *minimizing geodesic* if it is *essential* (homotopic to one of the boundaries), and its length is minimal among the essential curves.

Definition 2.1. A *zigzag* Z is a collection of homotopies $H_1, \ldots, H_n$ with the following properties:

1) $H_i(1) = H_{i+1}(0)$
2) H_i alternates between outward and inward monotone homotopies, i.e., each of the H_i is a monotone homotopy, but for any $i \in \{1, \ldots, n-1\}$, the concatenation of H_i and H_{i+1} is not.

We define $\gamma_0 = H_1(0)$ and $\gamma_i = H_i(1)$ for $1 \leq i \leq n$.

Each H_i goes from γ_{i-1} to γ_i. We define the *order* of Z, $ord(Z)$, to be n.

We will also need the following definitions and a theorem from the article of Chambers and Rotman [**9**].

Definition 2.2 ([**9**, Definition 0.6])**.** Let $\alpha : [0, 1] \longrightarrow M$ and $\beta : [0, 1] \longrightarrow M$ be two simple closed curves in a Riemannian manifold M. If every two points of intersection between α and β are consecutive on α if and only if they are consecutive on β, then α and β are said to satisfy the simple intersection property.

When α, β defined in 2.2 do not satisfy the simple intersection property, we will say that they are *meandering* with respect to each other.

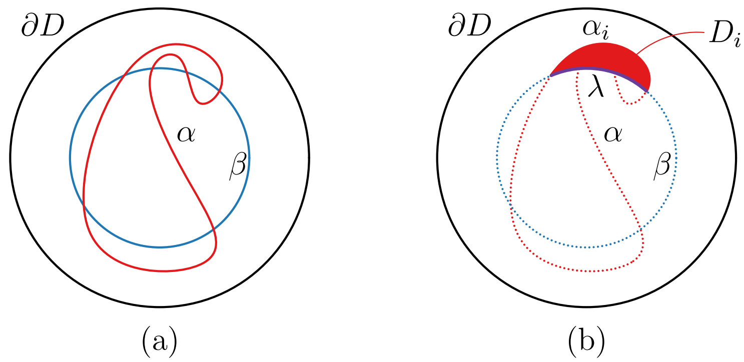

Figure 2. Meandering curves.

Definition 2.3. Let α and β be two simple closed curves in a closed topological 2-disc D. Let $\alpha_i = \alpha|_{[t_i,t_{i+1}]}$ be an arc of α, such that the interior of the arc does not intersect β, while its endpoints $\alpha(t_i), \alpha(t_{i+1}) \in \beta$. Then these points subdivide β into two arcs. Let λ be an arc that together with α_i bounds a disc in the closed annulus $A(\partial D, \beta(t))$ between ∂D and β. Then we will call λ a *corresponding arc*. We will refer to the disc D_i with the boundary $\alpha_i \cup \lambda$ as a *corresponding disc*. (See fig. 2 (b), where the disc that corresponds to arc α_i is shaded.) Note that α and β may intersect an infinite number of times; this definition still holds.

We first prove that any (non-monotone) homotopy can be approximated by a zigzag.

Proposition 2.4. *Suppose that there is a contraction of ∂D through curves of length less than L. Then there exists a zigzag of order n such that $\gamma_0 = \partial D$ and γ_n is a constant curve, and such that all curves of all homotopies have length less than L.*

Proof. First, by a result of Chambers and Liokumovich [7, Theorem 1.1], we know that there exists a contraction of ∂D through *simple* closed curves of length less than L.

We say that a corresponding disc between two arcs α and α' is δ-thin if the homotopic Fréchet distance between the two curves is less than δ, where the homotopic Fréchet distance between two curves is defined by considering all homotopies $H : [0,1] \times [0,1] \to D$ from $H(\cdot, 0) = \alpha$ to $H(\cdot, 1) = \alpha'$ (up to reparametrizations), and taking $\inf_H \sup_{s\in[0,1]} \text{length}(H(s,\cdot))$. In other words, for each homotopy from α to α' (allowing reparametrizations), we consider the length of the longest curve $H(s,\cdot)$ in that homotopy; taking the infimum of this quantity over all of these homotopies yields the homotopic Fréchet distance [4]. Similarly, an annulus $A(\alpha, \beta)$ is δ-thin if the two boundary curves have homotopic Fréchet distance less than δ. Now, we consider a discretized contraction H; we consider an increasing sequence of n values $t_1, \dots, t_n$ in $[0,1]$ so that

- $t_0 = 0$ and $H(0) = \partial D$,
- H on $[t_i, t_{i+1}]$ is a homotopy through curves of length less than L,
- $t_n = 1$ and $H(1)$ is a constant curve, and
- for $0 \le i \le n-1$, if $H(t_i)$ and $H(t_{i+1})$ intersect, they have the simple intersection property and the corresponding discs are δ-thin, for δ to be determined later. If they do not intersect, the annulus $A(H(t_i), H(t_{i+1}))$ is δ-thin.

We begin with the original contraction $\tilde{H}$ produced by the theorem of Chambers and Liokumovich. Without loss of generality, we may assume that the only constant curve in the contraction occurs at $t = 1$. If this is not the case, then we choose t_0 to be the smallest element of $[0,1]$ such that $\tilde{H}(t_0)$ is a constant curve, and we apply the rest of the argument to the contraction formed by restricting $\tilde{H}$ to $[0, t_0]$.

Next, we select $t^* \in (0,1)$ sufficiently large so that $\tilde{H}(t^*)$ can be contracted in $D(\tilde{H}(t^*))$ through disjoint closed curves of length less than L, which are all simple except for the final constant curve. Let this contraction of $\tilde{H}(t^*)$ be denoted by $K : [t^*, 1] \times S^1 \to D$.

Since H is smooth and the interval $[0, t^*]$ is compact, there is a $\delta > 0$ and an $\varepsilon > 0$ such that, for every $s_1, s_2 \in [0, t^*]$, if $|s_1 - s_2| < \varepsilon$, then $\tilde{H}(s_2)$ is contained in the δ-tubular neighborhood of $\tilde{H}(s_1)$. Furthermore, there are parametrizations γ_1 and γ_2 of $\tilde{H}(s_1)$ and $\tilde{H}(s_2)$, respectfully, such that $\gamma_2(x) = \gamma_1(x) + f(x)v(x)$, where $v(x)$ is the outward unit vector to γ_1, and $f(x)$ is a real number with $|f(x)| < \delta$ (note that outward is with respect to $D(\gamma_1)$). Due to this property, if $\tilde{H}(s_1)$ and $\tilde{H}(s_2)$ do not intersect, then $A(\tilde{H}(s_1), \tilde{H}(s_2))$ is δ-thin, and if they do intersect, then they have the simple intersection property, and the corresponding discs are all δ-thin.

To complete the proof, let $n = \lceil \frac{t^*}{\varepsilon} \rceil$, and take our discretized sequence to be $t_i = i\frac{t^*}{n}$ for $i \in \{0, \dots, n\}$, and the contraction H is defined as $H = \tilde{H}$ on $[t_i, t_{i+1}]$ for all $i \in \{0, \dots, n-1\}$. Since K is monotone, we can find a sufficiently large positive integer m so that setting $t_i = t^* + (i-n)\frac{1-t^*}{m}$ for $i \in \{n, \dots, n+m\}$, and setting $H = K$ on $[t_i, t_{i+1}]$ for $i \in \{n, \dots, n+m-1\}$ completes the proof.

Now, if $H(t_i)$ and $H(t_{i-1})$ intersect, for each $0 < i < n$, we define an auxiliary curve $H(t_i)^f$ from $H(t_i)$: $H(t_i)^f$ is obtained from $H(t_i)$ by considering all of the arcs of $H(t_i)$ in $D(H(t_{i-1}))$ and replacing the other ones by the arcs they correspond to in $H(t_{i-1})$. Then we claim that there are monotone homotopies between $H(t_i)$ and $H(t_i)^f$, and between $H(t_i)^f$ and $H(t_{i+1})$ such that the intermediate curves have length less than L. Indeed, one can go from one to the other using monotone homotopies that interpolate within the corresponding discs,

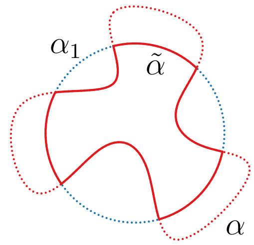

Figure 3. Construction of $\tilde{\alpha}$.

and if δ is chosen small enough, this interpolation can be done with an arbitrarily small overhead on the lengths of the curves. If $H(t_i)$ and $H(t_{i-1})$ do not intersect, for δ small enough, the δ-thin assumption implies that there exists a monotone homotopy between $H(t_{i-1})$ and $H(t_i)$, such that the intermediate closed curves have length less than L. Gluing together all of these monotone homotopies, we obtain a zigzag with curves of length at most L. q.e.d.

One of our main technical tools is a technique to modify a monotone homotopy when it crosses a minimizing geodesic. The general setting is when we have a monotone homotopy H between two curves α_0 and α_1, and α is a third simple closed curve that is a minimizing geodesic in $A(\alpha, \alpha_0)$. Then we can use α to "shortcut" the homotopy H. Depending on the relative positions of α_1, there are three variants of this shortcutting argument, leading to three different outcomes: they are summarized in the following proposition.

Proposition 2.5. *Let H be a monotone homotopy between simple closed curves α_0 and α_1 such that the intermediate curves have length less than L and let α be another simple closed curve, disjoint from α_0 and such that α is a minimizing geodesic in $A(\alpha_0, \alpha)$. Then:*

1) *If α is entirely contained and essential in $A(\alpha_0, \alpha_1)$, then there exists a monotone homotopy between α and α_1 where the intermediate curves have length less than L.*
2) *If α is entirely contained and non-essential in $A(\alpha_0, \alpha_1)$, then there exists a monotone homotopy between α and a constant curve p where p is a point on α, and where the intermediate curves have length less than L.*
3) *If α has the simple intersection property with α_1, then if we denote by $\tilde{\alpha}$ the curve obtained from α_1 by replacing segments of α_1 in $A(\alpha_0, \alpha)$ with the corresponding arcs (see Figure 3), there exists a monotone homotopy between α and $\tilde{\alpha}$ where the intermediate curves have length less than L.*

Although not explicitly stated in Chambers and Rotman [9], the first case of this Proposition is implicit in the proof of their Theorem 0.7.

More precisely, their proof is divided in two steps, and this is the result obtained in Step 1. The proof of the other two cases of Proposition 2.5 also follows closely the arguments of the proof of Theorem 0.7. For the sake of completeness, we include the full proof below.

Proof. The general idea is that in all three settings one can take each intermediate curve of the homotopy H and replace the portions that are outside of the target annulus using α. More precisely, if we denote by $(\alpha_t)_{t\in[0,1]}$ the curves of the homotopy H, we do the following in each case:

1) Let t_0 denote the first time t that α_t intersects α. Then for all $t \geq t_0$, we replace in each curve α_t the segments that are in the interior of the annulus $A(\alpha_0, \alpha)$ with their corresponding arcs and consider the family A of the closed curves $(\alpha_t)_{t\in[t_0,1]}$.
2) Let t_0 and t_1 denote respectively the first and last time t that α_t intersects α. Then for all $t_0 \leq t \leq t_1$, we replace in each curve α_t the segments that are in the interior of the annulus $A(\alpha_0, \alpha)$ with their corresponding arcs. Note that α_{t_1} is a contractible curve of α (viewed as a set), and can thus be contracted to a point $p \in \alpha$ while monotonically decreasing its length. This contraction is realized through curves $(\alpha_t)_{t\in[t_1,1]}$. We now consider the family A of closed curves $(\alpha_t)_{t\in[t_0,1]}$.
3) Let t_0 denote the first time t that α_t intersects α. Then for all $t \geq t_0$, we replace in each curve α_t the segments that are in the interior of the annulus $A(\alpha_0, \alpha)$ with their corresponding arcs and consider the family A of the closed curves $(\alpha_t)_{t\in[t_0,1]}$. Note that the new α_1 coincides with $\tilde{\alpha}$.

Now, the rest of the proof is the same in all three cases. Since α is minimizing in $A(\alpha_0, \alpha)$, the corresponding arcs are always shorter than the arcs that they replace. Thus, the families A that we obtain contain intermediate curves of length less than L. Furthermore, α_{t_0} and α_1 are the starting and ending curves of the target homotopy in all three cases. However, the families A fail to be monotone homotopies because they are neither homotopies (there can be discontinuities) nor monotone (the curves are not even simple). The first issue is solved by interpolation and the second one by perturbation.

Discontinuities only appear at times t when the intersection between α_t and α is not transversal. Figure 4 depicts such a situation. Here α_{t_2} touches α at point Q. There are two ways to replace the segments of α_{t_2} in the neighborhood of Q, (see Figure 5 (a) that depicts this situation locally). One way is to replace the segment of α_{t_2} that connects the points Q_1 and Q_2 that lies in the annulus $A(\alpha_0, \alpha)$ by the path P_1, (see Figure 5 (b)). Let us call this replacement the type 1 replacement. Another way is depicted in Figure 5 (c). Here we replace the segment

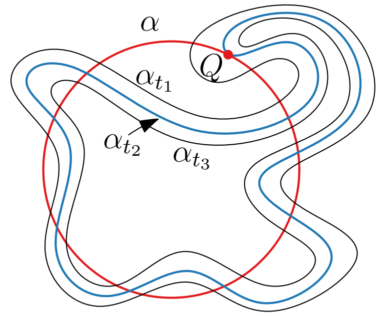

Figure 4. Discontinuities might occur. . . .

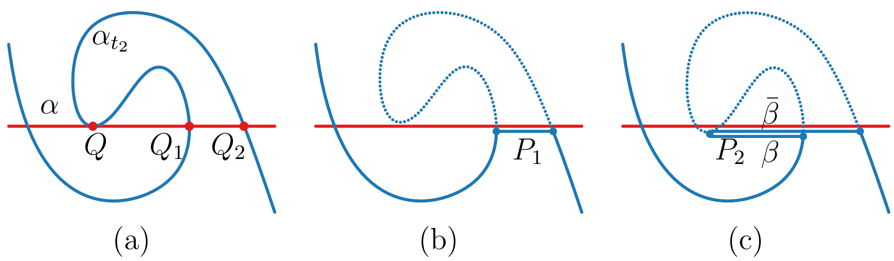

Figure 5. . . . but they can be fixed by interpolating with a path homotopy.

of α_{t_2} that connects Q_1 and Q_2 by P_2. P_2 is a path that consists of two paths: the first one replaces the segment of α_{t_2} that connects Q and Q_2, while the second one, β, replaces the segment of α_{t_2} that connects Q_1 and Q. Let us call this replacement the type 2 replacement. Since our procedure only replaces segments in the *interior* of the annulus $A(\alpha_0, \alpha)$ with their corresponding arcs, it always chooses the type 2 replacement.

Now, there is only one way to replace the relevant part of α_{t_1}, with a curve that is close to α_{t_2} and is contained in $A(\alpha_0, \alpha_{t_2})$ (Figure 4). If we want the procedure to result in a homotopy, this fits well with our choice of type 2 replacement on α_{t_2}. On the other hand, there is also only one type of replacement that can be performed on α_{t_3}, with a curve that is close to α_{t_2} and is contained in $A(\alpha_{t_2}, \alpha_1)$. And as t_3 goes to t_2, this converges into a type 1 replacement for α_{t_2}. Hence, we have a discontinuity at t_2. To avoid this discontinuity, note that $P_2 = \beta * \bar{\beta} * P_1$ (see Figure 5 (c)). Here, $\bar{\beta}$ denotes the path β traversed in the opposite direction, and $a{*}b$ denotes the curve formed by concatenating the curves a and b. Therefore, P_1 and P_2 can be connected by the obvious length non-increasing path homotopy, which amounts to contracting $\beta{*}\bar{\beta}$ to Q_1. This path homotopy extends to a homotopy between the two curves derived from the two replacement choices for α_{t_2}. Thus, including the homotopy between the two different resulting curves corresponding to type 1 and type 2 replacements solves the discontinuity problem at time t_2. Doing so for each time t when the intersection between α_t and α is not transverse makes A into a homotopy between α_{t_0} and α_1.

Finally, observe that while the curves in the homotopy A may not be simple, they do not feature transversal intersections, since the shortcutting procedure replaces all the segments outside of the annulus $A(\alpha_0, \alpha)$ by their corresponding arcs. Furthermore, since H was a monotone homotopy, there was no transverse intersection between α_t and $\alpha_{t'}$ for $t \neq t'$ before the replacement procedure, and thus by the same argument there are none afterwards either. Now, by applying an arbitrarily slight perturbation to all of the curves in A in a continuous way, we can make all the curves simple while still having no transverse intersection pairwise and all having length less than L. This yields a monotone homotopy and concludes the proof. q.e.d.

The proof of Theorem 1.2 relies on Propositions 2.6 and 2.8, which allow us to modify small portions of zigzags. The first one follows rather directly from Proposition 2.5, but the second one requires more work. The proof of Theorem 1.3 relies on Proposition 2.8 and a small variant of Proposition 2.6, which is stated in Proposition 2.9.

Proposition 2.6. *Suppose that Z is an order 2 zigzag through curves of length less than L. If γ_1 is not a minimizing geodesic in $A(\gamma_1, \gamma_2)$, and if a minimizing geodesic γ in this annulus also lies in the interior of $A(\gamma_0, \gamma_1)$ and is essential in it, then there is a zigzag Z' where the intermediate curves have length less than L and such that*

1) *$ord(Z') = 2$.*
2) *γ'_1 minimizes in $A(\gamma'_1, \gamma'_2)$.*
3) *$\gamma'_0 = \gamma_0$, $\gamma'_1 = \gamma$, and $\gamma'_2 = \gamma_2$.*

Suppose that Z is an order 2 zigzag where the intermediate curves have length less than L, and that γ_0 is a minimizing geodesic in $A(\gamma_1, \gamma_2)$, or that γ_2 is a minimizing geodesic in $A(\gamma_0, \gamma_1)$. Then there is an order 1 zigzag Z' (i.e., a monotone homotopy) through curves of length less than L and such that $\gamma'_0 = \gamma_0$, and $\gamma'_1 = \gamma_2$.

Proof. The first part of the proposition follows from two applications of Proposition 2.5(1). We first apply it to the reversal of the homotopy H_1 and the curve γ, and then to the homotopy H_2 and the curve γ. This yields two new homotopies H'_0 and H'_1, going respectively from γ_0 to γ and from γ to γ_2; their concatenation satisfies the needed properties.

For the second part of the proposition, let us first deal with the first case where γ_0 is a minimizing geodesic in $A(\gamma_1, \gamma_2)$. Then one application of Proposition 2.5(1) to the homotopy H_2 and γ_0 yields the homotopy from γ_0 to γ_2. The other case is obtained by applying the theorem to the reversal of H_1 and γ_2 instead. q.e.d.

Lemma 2.7. *Suppose that Z is a zigzag of order 2, and that γ_0 is a minimizer in $A(\gamma_0, \gamma_1)$. Then there exists an essential curve γ which is a minimizer in $A(\gamma_1, \gamma_2)$, and which has the simple intersection property with γ_0.*

Proof. We begin by choosing an essential minimizing curve α in $A(\gamma_1, \gamma_2)$. Let ϱ be a segment of γ_0 whose endpoints are intersections between α and γ_0, and whose interior is contained in the interior of $A(\gamma_1, \alpha)$. Let the endpoints of ϱ be $\varrho(0)$ and $\varrho(1)$.

From α and ϱ we can define two auxiliary curves: one which goes from $\varrho(0)$ to $\varrho(1)$ following α and then back to $\varrho(0)$ along ϱ, and one which goes from $\varrho(1)$ to $\varrho(0)$ following α and then back to $\varrho(1)$ along ϱ. Let the first curve be β_1, and let the other one be β_2. Note that both β_1 and β_2 are contained in $A(\gamma_1, \alpha)$ and are simple closed curves. Thus, in this annulus, exactly one of β_1 or β_2 is essential; without loss of generality, we may assume that it is β_1. Since β_2 is not essential, it bounds a disc within $A(\gamma_1, \alpha)$ which we call by a slight abuse of language its *interior*. If ϱ lies in the boundary of a portion of the interior of β_2 outside of $A(\gamma_0, \gamma_1)$, as in the middle picture of Figure 6 then we do nothing.

If ϱ lies on the boundary of a portion of the interior of β_2 inside of $A(\gamma_0, \gamma_1)$, as in the right picture of Figure 6, we claim that β_1 has total length not greater than that of α. Indeed, let us first build an auxiliary closed curve in the following way: take all the segments Σ of β_2 that lie in the interior of $A(\gamma_0, \gamma_1)$ and have their endpoints on γ_0. The segments in Σ are also segments of α, and since γ_0 is a minimizing geodesic in $A(\gamma_0, \gamma_1)$, any $\sigma \in \Sigma$ is at least as long as its corresponding arc on γ_0. The closed curve α' is obtained by replacing all the segments in Σ from α by their corresponding arcs on γ_0; it is not longer than α. Now, α' may not be simple, in particular there may be double points on $\varrho(0)$ or $\varrho(1)$. Such a double point $\alpha'(t_1) = \alpha'(t_2)$ for $t_1 \neq t_2$ cuts α' into two subcurves, one of which, say $\alpha'_{|[t_1,t_2]}$, is contractible in $A(\gamma_1, \gamma_2)$. One can shortcut α' even more by removing such a contractible portion, i.e., replacing α' by the closed subcurve $\alpha'_{|[t_2,t_1]}$. After removing all these contractible subcurves, we obtain the curve β_1, which is by construction not longer than α' and thus not longer than α.

We now build γ in the following way. Since γ_0 has bounded length, there are countably many segments of γ_0 which satisfy the above properties. Let these segments be $\varrho_1, \varrho_2, \ldots$; we apply the above procedure to $\alpha = \omega_0$ and ϱ_1 to obtain a curve ω_1. If a segment of γ_0 satisfies the above properties with respect to ω_1, then it also satisfies those properties with respect to ω_0 $(= \alpha)$. If ϱ_2 is still one of these segments, then we apply the procedure to ω_1 and ϱ_2 to form ω_2. We continue to do this for all ϱ_i to form a sequence of curves $\omega_1, \omega_2, \ldots$. All of these curves are minimizers in $A(\gamma_1, \gamma_2)$, and so all lie in this annulus, all are smooth, and all have length bounded by L. By the Arzelà-Ascoli theorem, there

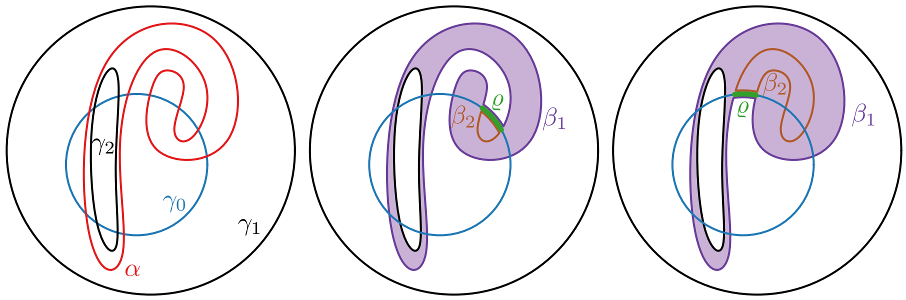

Figure 6. The different curves in the proof of Lemma 2.7. For the ϱ in the middle diagram, we do nothing, while for the ϱ in the right diagram we can shortcut α by replacing it with β_1.

is a curve to which these curves converge; we let this curve be γ. γ is still a minimizer in $A(\gamma_1, \gamma_2)$, and γ has the simple intersection property with γ_0. If it did not have the simple intersection property with γ_0, then there is a segment ϱ of γ_0 whose endpoints are also in γ, whose interior is contained in the interior of $A(\gamma_1, \gamma)$, and which produces a β_1 and a β_2 which are in the second case. However, then $\varrho = \varrho_i \in \{\varrho_1, \varrho_2, \dots\}$, and so $\varrho = \varrho_i$ would not satisfy the above properties with respect to ω_i, and so it would also not satisfy the above properties with respect to γ, yielding a contradiction. q.e.d.

Proposition 2.8. *Suppose that Z is a zigzag of order 3 where the intermediate curves have length at most L and such that γ_1 is a minimizing geodesic in $A(\gamma_1, \gamma_2)$, but is not a constant curve.*

Then one of the following two cases is true:

Case a. *There is a zigzag Z' of order 3 such that*

1) $\gamma'_0 = \gamma_0$, $\gamma'_3 = \gamma_3$, *and* $\gamma'_2 = \gamma_2$.
2) *There exists a minimizing geodesic $\gamma \in A(\gamma'_2, \gamma'_3)$ which is fully contained in the interior of $A(\gamma'_1, \gamma'_2)$.*
3) *γ'_1 is a minimizing geodesic in $A(\gamma'_1, \gamma'_2)$.*
4) *All curves in Z' have length less than L.*

Case b. *There exists a zigzag Z' of order 1 such that*

1) $\gamma'_0 = \gamma_0$.
2) *γ'_1 is a constant curve.*
3) *All curves in Z' have length less than L.*

Proof. We begin by applying Lemma 2.7 to the order 2 zigzag from γ_1 to γ_2 to γ_3 to obtain an essential minimizing geodesic γ in $A(\gamma_2, \gamma_3)$ which has the simple intersection property with γ_1.

If γ lies in $A(\gamma_1, \gamma_2)$, and is essential in this annulus, then we are done as Case a is satisfied.

We now divide the remainder of the proof into two cases:

(i) If γ is not entirely contained in $A(\gamma_1, \gamma_2)$, we will show that case a. holds.
(ii) If γ is contained in $A(\gamma_1, \gamma_2)$, and is non-essential there, we will show that case b. holds.

Case (i) Suppose that γ is not entirely contained in $A(\gamma_1, \gamma_2)$. Note that one can always modify the homotopy H_3 to obtain a new monotone homotopy H_3' between γ_2 and γ_3, such that the lengths of the curves in H_3' are less than L and γ is one of the curves of H_3'. One achieves this by applying Proposition 2.5(1) to the reversal of the homotopy H_3 and γ, and then to H_3 and γ and concatenating the two resulting homotopies. Thus, without loss of generality, we assume that $\gamma = (H_3)_t$ for some $t \in [0, 1]$.

Since γ is not entirely contained in $A(\gamma_1, \gamma_2)$, the new zigzag Z' is obtained in the following manner. We define a new curve $\tilde{\gamma}$ by replacing the segments of γ in $A(\gamma_1, \gamma_2)$ with the corresponding segments of γ_1. Then, using Proposition 2.5(3) on H_3 and γ_1, we form an auxiliary monotone homotopy $\tilde{H}$ from γ_1 to $\tilde{\gamma}$.

To form our new homotopy, we append the homotopy $\tilde{H}$ to the end of H_1, and we append $\tilde{H}$ in reverse direction to the beginning of H_2. Clearly, properties 1 and 4 are satisfied.

Property 2 follows from the fact that γ and γ_1 have the simple intersection property.

To prove property 3, we fix any essential curve γ' in $A(\gamma_1', \gamma_2')$. If we replace segments of γ' which lie in $A(\gamma_1, \gamma_2)$ with segments of γ_1, then replace segments of the resulting curve which lie in $A(\gamma_2', \gamma)$ with segments of γ, we obtain $\tilde{\gamma} = \gamma_1'$. This procedure does not increase the length, so this shows that the length of γ' is greater than or equal to the length of γ_1', completing the proof.

Case (ii) Suppose that γ is contained in $A(\gamma_1, \gamma_2)$, and is non-essential in it. In this case, the disc bounded by γ does not intersect the disc bounded by γ_1. Thus, by monotonicity, the homotopy H_3 "sweeps" $D(\gamma_1)$ completely. Thus we are in the situation to apply case (2) of Proposition 2.5 to the homotopy H_3 and γ_1. This yields a homotopy $\tilde{H}$ between γ_1 and a point p on γ_1. We then concatenate $\tilde{H}$ to the end of H_1 to form H_1'. Clearly, both properties are satisfied. q.e.d.

Proposition 2.9. *Suppose that Z is a zigzag of order 2 on a Riemannian sphere such that all curves have length less than L, and such that γ_0 is a constant curve, but γ_1 and γ_2 are not constant curves. Furthermore, assume that the orientation of the sphere is such that the discs bounded by curves close to γ_0 in the first monotone homotopy are close to the image of γ_0. Then there exists an order 2 zigzag Z' where γ_0' is a constant curve in $D(\gamma_1)$, $\gamma_2' = \gamma_2$, and γ_1' is a minimizing geodesic in $A(\gamma_1', \gamma_2')$.*

Proof. Let γ be a minimizing geodesic in $A(\gamma_1, \gamma_2)$. There are two possibilities:

1) γ_0 is contained in both $D(\gamma_1)$ and $D(\gamma)$.

2) γ_0 is contained in exactly one of $D(\gamma_1)$ and $D(\gamma)$.

If the first condition is true, then the conclusion follows from Proposition 2.6. If the second condition is true, then H_1 "sweeps" $D(\gamma)$, and we apply case (2) of Proposition 2.5 to the reversal of H_1 and γ to obtain a monotone homotopy from γ to a point. Furthermore, applying case (1) of Proposition 2.5 to H_2 and γ yields a monotone homotopy from γ to γ_2. Concatenating these gives the result. q.e.d.

3. Proof of Theorems 1.2 and 1.3

We first find a zigzag which starts at the boundary of the Riemannian disc, ends at a constant curve, and traverses curves of length less than L which minimizes the order of the zigzag.

Proposition 3.1. *Suppose that there exists a contraction of ∂D through curves of length less than L. Then there is a zigzag Z of finite order, which consists of curves of length less than L, and which begins on ∂D, and ends at a point. Furthermore, for every zigzag $\tilde{Z}$ with these properties, the order of $\tilde{Z}$ is greater than or equal to the order of Z.*

The proof follows directly from Proposition 2.4, and from the fact that the order of a zigzag is a positive integer. We will need one more lemma before we can prove our two theorems.

Lemma 3.2. *Suppose that Z is a zigzag of order $n \geq 2$ through curves of length less than L, and suppose that at most the initial and final curves are constant. Suppose further that γ_1 is a minimizing geodesic in $A(\gamma_1, \gamma_2)$. Then one of the following is true. First, there exists a zigzag Z' of order n with $\gamma'_0 = \gamma_0$ and $\gamma'_n = \gamma_n$, every γ_i is a minimizing geodesic in $A(\gamma_i, \gamma_{i+1})$ for all $i \in \{1, \ldots, n-1\}$, and some minimizing geodesic in $A(\gamma_{i+1}, \gamma_{i+2})$ is contained and essential in $A(\gamma_i, \gamma_{i+1})$ for all $i \in \{1, \ldots, n-2\}$. Second, there exist two zigzags, Z'_1 and Z'_2 of orders $m_1 > 0$ and $m_2 > 0$ with $m_1 + m_2 = n$ through curves of length less than L, and such that the first curve of Z'_1 is equal to the first curve of Z, the last curve of Z'_2 is equal to the last curve of Z, the last curve of Z'_1 is a constant curve, and the first curve of Z'_2 is equal to the same constant curve.*

Proof. We will prove this lemma by induction on n, and by using Proposition 2.8 and Proposition 2.6. If $n = 2$, there is nothing to prove. If $n = 3$, then we apply Proposition 2.8 to Z, followed by applying Proposition 2.6 to the final order 2 zigzag. If, during the process, we produce a zigzag of smaller order which ends at a constant curve, then we terminate this procedure.

For the inductive step, we first apply the induction hypothesis to the order $n-1$ zigzag at the beginning of Z. If we obtain a zigzag of smaller order which ends at a constant curve, then we are in the second case of the lemma and we are done. Otherwise, we obtain a new zigzag Z' where γ_{n-2} is a minimizing geodesic in $A(\gamma_{n-2}, \gamma_{n-1})$. Thus we are in a position to apply Proposition 2.8 to the order 3 zigzag at the end of Z', from γ_{n-3} to γ_n. If we are in case b of that proposition, we are done since we obtain a zigzag of smaller order ending at a constant curve.

Otherwise, we obtain a new zigzag, which we replace into Z' to yield Z''. Since γ_{n-2} may have moved in the last step, it may be the case that γ_{n-3} is not a minimizing geodesic in $A(\gamma_{n-3}, \gamma_{n-2})$ anymore. In order to fix this, we apply the induction hypothesis once again, this time to the $n-2$ zigzag at the beginning of Z'', yielding yet another zigzag Z''' (once again, we are done if we are in the second case of the lemma). Since γ_{n-2} and γ_{n-1} have not been changed in this last step, we still have that γ_{n-2} is a minimizing geodesic in $A(\gamma_{n-2}, \gamma_{n-1})$, and some minimizing geodesic in $A(\gamma_{n-1}, \gamma_n)$ is contained and essential in $A(\gamma_{n-2}, \gamma_{n-1})$. Now, either γ_{n-1} is a minimizing geodesic in $A(\gamma_{n-1}, \gamma_n)$ and we are done, or we can apply Proposition 2.6 to the final order 2 zigzag of Z'''. The resulting zigzag fulfills all the properties of the lemma. q.e.d.

We now have all the tools to prove our main theorems.

Proof of Theorem 1.2. Let Z be a zigzag which satisfies the conclusion of Proposition 3.1. If the order of Z is equal to 1, then the proof is finished. As such, assume that $ord(Z) > 1$. Since the zigzag must end at a constant curve, $ord(Z) \geq 3$. We may further assume that no other curve in the zigzag is a constant curve. Additionally, since $\gamma_0 = \partial D$, $A(\gamma_1, \gamma_2)$ is contained in $A(\gamma_0, \gamma_1)$. As a result, we can apply Proposition 2.6 to replace Z with a zigzag with the property that γ_1 is minimizing in $A(\gamma_1, \gamma_2)$ (this also uses the fact that we cannot produce a zigzag of shorter order from ∂D to a constant curve). As a result, we may assume that Z has this property.

We now apply Lemma 3.2 to Z. Since we cannot find a zigzag of smaller order which begins at ∂D and ends at a constant curve, the result is an order n zigzag satisfying the conclusions of the lemma. In particular, γ_{n-1} must be a minimizing geodesic in $A(\gamma_{n-1}, \gamma_n)$, but must not be a constant curve. However, γ_n is a constant curve, and so γ_{n-1} must also be constant, having length 0. This is a contradiction, completing the proof. q.e.d.

We can use a very similar technique to prove Theorem 1.3:

Proof of Theorem 1.3. To prove Theorem 1.3, we proceed in a similar way. We first apply Theorem 1.2 from [**8**], which tells us that we can replace our sweepout f of our Riemannian sphere (S^2, g) of curves of length less than L by a sweepout which contains only simple closed curves and constant curves. Let this sweepout be parametrized by $[0, 1]$, where 0 and 1 are mapped to the same constant curve. Since it is smooth, we can find a finite number of subintervals $I_1, \ldots, I_k$ of $[0, 1]$ such that the boundaries of I_i are mapped to constant curves, the interior of I_i is mapped to simple closed curves, and the degree of the map of f restricted to I_i is $d_i \neq 0$. Furthermore, the sum of all of the degrees of these maps is equal to 1, the degree of the map. As a result, there is at least one such map that has odd degree.

We now apply Proposition 3.1 to this map to produce a zigzag which starts and ends at constant curves, and contains no other constant curves. Since this zigzag is homotopic to the original map, it also has odd degree. Let Z be a minimal zigzag of odd degree on the sphere which begins and ends at a constant curve, and only passes through simple closed curves, and which has minimal order.

If Z has order one, we are done. Otherwise, Z has order at least 3, and we first apply Proposition 2.9, and then Lemma 3.2 to Z; if we can divide it into two zigzags (the second possibility of the lemma), then the concatenation is homotopic to Z, and so has odd degree as a map, and so one of the zigzags must have odd degree as a map but order smaller than n, which contradicts the minimality of the order of Z. If we are in the first conclusion of Lemma 3.2, then the result is homotopic to Z, and so has odd degree as a map. As in the proof of Theorem 1.2, γ_{n-1} must be a constant curve, as it must be a minimizing geodesic in $A(\gamma_{n-1}, \gamma_n)$, and γ_n is a constant curve with length 0. Thus, the degree of the last segment of Z is 1, which contradicts the minimality of the order of Z. This completes the proof. q.e.d.

References

[1] Graham R. Brightwell and Peter Winkler, *Submodular percolation*, SIAM J. Disc. Math **23** (2009), no. 3, 1149–1178.

[2] Benjamin Burton, Erin Wolf Chambers, Marc van Kreveld, Wouter Meulemans, Tim Ophelders, and Bettina Speckmann, *Computing optimal homotopies over a spiked plane with polygonal boundary*, Proceedings of the 25th Annual European Symposium on Algorithms, 2017.

[3] Erin W. Chambers and David Letscher, *On the height of a homotopy*, Proceedings of the 21st Canadian Conference on Computational Geometry, 2009, pp. 103–106.

[4] Erin Wolf Chambers, Éric Colin de Verdière, Jeff Erickson, Sylvain Lazard, Francis Lazarus, and Shripad Thite, *Homotopic fréchet distance between curves or, walking your dog in the woods in polynomial time*, Computational Geometry **43** (2010), no. 3, 295–311, Special Issue on 24th Annual Symposium on Computational Geometry (SoCG'08).

[5] Erin Wolf Chambers, Arnaud de Mesmay, and Tim Ophelders, *On the complexity of optimal homotopies*, In preparation, 2017.

[6] Erin Wolf Chambers and Yusu Wang, *Measuring similarity between curves on 2-manifolds via homotopy area*, Proc. 29th Ann. Symp. on CG, ACM, 2013, pp. 425–434.

[7] Gregory R. Chambers and Yevgeny Liokumovich, *Converting homotopies to isotopies and dividing homotopies in half in an effective way*, Geometric and Functional Analysis **24** (2014), no. 4, 1080–1100.

[8] Gregory R. Chambers and Yevgeny Liokumovich, *Optimal sweepouts of a Riemannian 2-sphere*, Journal of the European Mathematical Society (2014), to appear.

[9] Gregory R. Chambers and Regina Rotman, *Monotone homotopies and contracting discs on riemannian surfaces*, Journal of Topology and Analysis (2016).

[10] Sariel Har-Peled, Amir Nayyeri, Mohammad Salavatipour, and Anastasios Sidiropoulos, *How to walk your dog in the mountains with no magic leash*, Discrete & Computational Geometry **55** (2016), no. 1, 39–73.

[11] Alexander Nabutovsky and Regina Rotman, *Linear bounds for lengths of geodesic loops on riemannian 2-spheres*, Journal of Differential Geometry **89** (2011), 217–232.

[12] Alexander Nabutovsky and Regina Rotman, *Length of geodesics and quantitative morse theory on loop spaces*, Geometric and Functional Analysis **23** (2013), 367–414.

[13] Alexander Nabutovsky and Regina Rotman, *Linear bounds for lengths of geodesic segments on riemannian 2-spheres*, Journal of Topology and Analysis **5** (2013), 409–438.

Department of Computer Science
Saint Louis University
220 N. Grand Ave.
Saint Louis, MO
USA
E-mail address: echambe5@slu.edu

Department of Mathematics
Rice University
6100 Main St.
Houston, TX
USA
E-mail address: gchambers@rice.edu

LIGM, CNRS
Univ. Gustave Eiffel, ESIEE Paris
F-77454 Marne-la-Vallée
France
E-mail address: arnaud.de-mesmay@univ-eiffel.fr

Department of Mathematics and Computer Science
TU Eindhoven
5612 AZ Eindhoven
The Netherlands
E-mail address: t.a.e.ophelders@tue.nl

Department of Mathematics
University of Toronto
40 St. George Street
Toronto, ON M5S 2E4
Canada
E-mail address: rina@math.toronto.edu