TOPOLOGICAL AND PHYSICAL LINK THEORY ARE DISTINCT

ALEXANDER COWARD AND JOEL HASS
TOPOLOGICAL AND PHYSICAL LINK THEORY ARE DISTINCT

ALEXANDER COWARD AND JOEL HASS

Physical knots and links are one-dimensional submanifolds of $\mathbb{R}^3$ with fixed length and thickness. We show that isotopy classes in this category can differ from those of classical knot and link theory. In particular we exhibit a Gordian split link, a two-component link that is split in the classical theory but cannot be split with a physical isotopy.

1. Introduction

The theory of knots and links studies one-dimensional submanifolds of $\mathbb{R}^3$. These are often described as loops of string, or rope, with their ends glued together. Real ropes however are not one-dimensional, but have a positive thickness and a finite length. Indeed, most physical applications of knot theory are related more closely to the theory of knots of fixed thickness and length than to classical knot theory. For example, biologists are interested in knotted curves of fixed thickness and length when studying properties of DNA [Cantarella et al. 1998] and protein molecules [Liang and Mislow 1994]. In these applications the thickness of the curve modeling the molecule plays an essential role in determining the possible configurations.

In this paper we show that the equivalence class of a link in $\mathbb{R}^3$ under an isotopy that preserves thickness and length can be distinct from the classical equivalence class under isotopy. We thus show for the first time that the theory of physically realistic curves of fixed thickness and length in $\mathbb{R}^3$ is distinct from the classical theory of knots and links.

The two most fundamental problems concerning physical knots and links are to show the existence of a Gordian unknot and a Gordian split link. A Gordian unknot is a loop of fixed thickness and length whose core is unknotted, but which cannot be deformed to a round circle by an isotopy fixing its length and thickness. A Gordian split link is a pair of loops of fixed thickness whose core curves can be split, or isotoped so that its two components are separated by a plane, but cannot be split by an isotopy fixing each component’s length and thickness. In this paper we establish the existence of such a link.

MSC2010: 53A04, 57M25.

Keywords: physical knot theory, Gordian, split link.
Theorem 1.1. A Gordian split link exists.

The proof of Theorem 1.1 is by explicit construction of a link, illustrated in Figure 1, that can be topologically but not physically split.

There has been extensive investigation into the properties of shortest representatives of physical knots, called tight [Cantarella et al. 1998] or ideal knots [Katritch et al. 1997], and into the possible existence of Gordian unknots [Buck and Simon 1997; Cantarella et al. 1998; Diao et al. 1999; Gonzalez and Maddocks 1999; Katritch et al. 1997]. A candidate Gordian unknot was suggested by Freedman, He and Wang [Freedman et al. 1994], who studied energies associated to curves in $\mathbb{R}^3$ and associated gradient flows [He 2002]. This curve was studied numerically by Pierański [1998], who developed a computer program called SONO (Shrink On No Overlaps) to numerically shorten a curve of fixed thickness while avoiding overlaps. The program unexpectedly succeeded in unraveling the Freedman–He–Wang example. However there are more complicated examples that do fail to unravel under SONO and hence give numerical evidence for the existence of Gordian unknots. Extensive tables of physical knots of minimal length in various isotopy classes have been experimentally derived [Ashton et al. 2011]. A proof of the existence of Gordian unknots or Gordian split links based on a rigorous analysis of such algorithms is plausible but has not yet been found. In related work, Nabutovsky [1995] showed that $n$-dimensional spheres of fixed thickness in $\mathbb{R}^{n+1}$ can be knotted for dimensions $n \geq 5$. Cantarella and Johnston [1998] showed that the theory of polygonal knots of fixed edge lengths is distinct from classical knot theory.

We now give precise definitions. We say that a knot or link $L$ in $\mathbb{R}^3$ is $r$-thick if it is differentiable and its open radius-$r$ normal disk bundle is embedded. This means that the collection of flat, radius-$r$ two-disks intersecting $L$ perpendicularly at their centers have mutually disjoint interiors. An isotopy of a knot or link maintaining $r$-thickness throughout is called an $r$-thick isotopy. By rescaling we may take $r = 1$ and take thick to mean 1-thick. A physical isotopy is a thick isotopy of a knot or link that preserves the length of each component. This reflects the real world properties of links composed of nonstretchable rope of fixed radius.
TOPOLOGICAL AND PHYSICAL LINK THEORY ARE DISTINCT

Theorem 1.1 is proved by an explicit construction of a two-component thick link \( L \) that is split but admits no physical isotopy splitting its components. To construct this link we begin by placing two points \( A \) and \( B \) at \((1, 0, 0)\) and \((-1, 0, 0)\). Let \( AB \) denote the straight line between these two points. The first component \( L_1 \) of \( L \) is any thick curve encircling \( AB \) in the \( xz \)-plane, disjoint from the open, radius-2 neighborhood of \( AB \). The length of \( L_1 \) is at least \( 4\pi + 4 \approx 16.566 \), and this length can be realized by taking \( L_1 \) to be the boundary of the radius-2 neighborhood of \( AB \) in the \( xz \)-plane. To construct the other component, join the two points \( A \) and \( B \) by an arc \( \alpha \) satisfying the following three conditions:

1. The union of \( \alpha \) with \( AB \) forms a nontrivial knot contained in the half-space \( y \geq 0 \).
2. The arc \( \alpha \) meets the \( xz \)-plane only at its endpoints and is perpendicular to the \( xz \)-plane at these points.
3. The union of \( L_1, \alpha \) and the reflection of \( \alpha \) across the \( xz \)-plane forms a thick link.

The union of \( \alpha \) and its reflection in the \( xz \)-plane is the second component \( L_2 \) of the thick link \( L \). Figure 1 shows an example of such a link.

Theorem 1.1 follows immediately from the following result, which gives an explicit lower bound on the length required for \( L_1 \), the unknotted component of \( L \), if \( L \) can be split by a physical isotopy.

**Theorem 1.2.** If there is a physical isotopy of \( L = L_1 \sqcup L_2 \) that splits its two components, then the length of \( L_1 \) must be at least \( 4\pi + 6 \approx 18.566 \).

Since the link \( L \) can be constructed with the length of the unknotted component \( L_1 \) equal to \( 4\pi + 4 \), this result implies Theorem 1.1.

The paper is arranged as follows. In Section 2 we give a lower bound on the boundary length of a disk of nonpositive curvature containing three disjoint disks of radius 1. In Section 3 we show that if a family of disks spanning \( L_1 \) gives a homotopy from a disk in the \( xz \)-plane to a disk disjoint from \( L_2 \) and each disk in the homotopy intersects a neighborhood of \( L_2 \) in at most two components containing points of \( L_2 \), then \( L_2 \) is unknotted. In Section 4 we bring these results together to prove Theorem 1.2. We conclude with a short list of open problems.

**2. An isoperimetric inequality**

To prove Theorem 1.2, we show that if the unknotted component \( L_1 \) has length less than \( 4\pi + 6 \), then there are severe restrictions on how the other component can pass through a natural spanning disc for \( L_1 \). This spanning disc, to be defined in Section 4, is a cone with cone angle at least \( 2\pi \), and hence a CAT(0) space. We are therefore led to finding a lower bound on the length of a curve in a CAT(0) surface.
Figure 2. The boundary of this disk has length at least $6 + 4\pi$.

that bounds a disk enclosing three or more nonoverlapping subdisks of radius 1, as in Figure 2. This next result is based on an argument for flat metrics given in [Cantarella et al. 2002].

**Proposition 2.1.** Let $P$ be a complete CAT(0) surface and let $D_1, D_2, D_3$ be three subdisks of $P$ with disjoint interiors and radius 1. Let $D \subset P$ be a disk containing $D_1, D_2, D_3$ such that $\partial D$ has distance at least 1 to any of $D_1, D_2, D_3$. Then the length of $\partial D$ is at least $4\pi + 6$.

**Proof.** Let $T$ be the triangle with vertices at the center points $c_1, c_2, c_3$ of the disks $D_1, D_2, D_3$. Each edge of $T$ has length at least 2. Let $a, b, c, d, e, f$ denote perpendicular rays from the sides of $T$ at its vertices, as in Figure 2. The curve $\partial D$ intersects each of $a, b, c, d, e, f$ in at least one point. We pick one such point for each, ordered cyclically around $\partial D$, and refer to the intervening arcs of $\partial D$ as the parts of $\partial D$ between them. Since the edge of $T$ between $c_1$ and $c_2$ is perpendicular to $a$ and to $b$, it realizes the minimal distance of any path between them. Thus the length of the part of $\partial D$ between $a$ and $b$ is at least 2. We can argue similarly for the length of $\partial D$ between $c$ and $d$, and between $e$ and $f$. Thus these three parts of $\partial D$ have total length at least 6.

The sum of the interior angles of $T$ is at most $\pi$, so the sum of the three angles between $f$ and $a$, between $b$ and $c$, and between $d$ and $e$ is at least $2\pi$. Radial projection projects the remaining parts of $\partial D$ onto three circular arcs with total angle at least $2\pi$. In a CAT(0) space, radius-decreasing radial projection onto a circle of constant radius is length-decreasing. Since a circle of radius 2 has length at least $4\pi$, it follows that the length of $\partial D$ is at least $6 + 4\pi$. □

**Remark.** A similar argument shows that a curve enclosing two disks has length at least $4\pi + 4$ and that the length of a curve enclosing $n > 3$ disks is at least $4\pi + 2n$ if the centers of the disks form the vertices of a convex polygon.
3. Sweepouts of solid tori

The following proposition gives a generalization of the fact that a 1-bridge knot is unknotted. It considers a generic 1-parameter family of disks, possibly singular, whose interiors sweep across a region containing a solid torus and concludes that if each disk meets the solid torus in at most two components that cross its core, then the solid torus is unknotted. To simplify the argument we restrict to a setting where $T$ and $c$ have a reflectional symmetry. Roughly speaking, this allows us to argue that if a family of arcs forms a partial spanning disk that fills in half of a curve, then the entire curve bounds a disk formed from reflection of this partial spanning disk and is thus unknotted.

**Proposition 3.1.** Let $T$ be a solid torus in $\mathbb{R}^3$ with core $c$, such that both $T$ and $c$ are symmetric under reflection $r$ in the xz-plane. Suppose there is a homotopy of the disc $g_t : D \to \mathbb{R}^3$, $t \in [-1, 1]$, with the following properties:

1. The curve $g_t(\partial D)$ is disjoint from $T$ for all $t \in [-1, 1]$.
2. The family of disks $g_t(D)$ is symmetric under reflection $r$ in the xz-plane; i.e., $g_0(D)$ is contained in the xz-plane and $g_{-t} = r \circ g_t$.
3. The preimage $g_0^{-1}(T)$ has two components, each containing a single point of $g_0^{-1}(c)$.
4. The disk $g_1(D)$ is disjoint from $c$.
5. For all $t \in [-1, 1]$ the preimage $g_t^{-1}(T) \subset D$ has at most two components that contain a point of $g_t^{-1}(c)$.
6. The map $g_t$ is generic with respect to the pair $(T, c)$.

Then $c$ is unknotted.

Assumption (6) means that $g_t$ is transverse to $c$ with the exception of a finite number of times $t$ at which a birth or death of a pair of points of $g_t^{-1}(c)$ occurs and that $g_t$ is transverse to $\partial T$ at these times. Additionally, $g_t$ is transverse to $\partial T$ except for a finite number of times at which $g_t^{-1}(\partial T)$ consists of finitely many simple closed curves and a single component that is either a bouquet of finitely many circles (at a general saddle-type singularity) or a point (at a birth or death singularity).

**Proof of Proposition 3.1.** To show that $c$ is unknotted, we will form a spanning disk $E$ for $c$ that is traced by a continuous family of arcs in $\mathbb{R}^3$, each arc having endpoints on $c$ and interior disjoint from $c$. These arcs are of two types. The first type will lie on $g_t(D)$ and vary continuously with $t$ except at finitely many times $t$ when it jumps from one arc on $g_t(D)$ to another; the second type will interpolate continuously between the arcs just before and just after these jumps.
For a map $g : D \to \mathbb{R}^3$ we call a point of $g^{-1}(c) \subset D$ a dot and a component of $g^{-1}(T)$ that contains at least one point of $g^{-1}(c)$ a dotted component. Thus $g_0^{-1}(T)$ contains two dotted components, each with a single dot. A birth or death changes the number of points of $g_t^{-1}(c)$ in a component of $g_t^{-1}(T)$ by two, oppositely oriented, as illustrated in Figure 3.

At time $t = 0$, the preimage $g_0^{-1}(c) \subset D$ consists of a pair of points, one in each dotted component. Let $\alpha_0$ be an arc joining these two points in $D$, with interior disjoint from $g_0^{-1}(c) \subset D$. As $t$ increases, we take $\alpha_t$ to vary continuously through arcs in $D$, joining dots in distinct dotted components with interiors disjoint from the dots. There is no obstruction to this while the collection of dots in $D$ is changing by an isotopy. As long as the number of dotted components does not drop, there are two possible obstructions to the extension of $\alpha_t$ as $t$ increases:

1. Part of the arc $\alpha_t$ may run between two dots that come together and disappear in a death singularity.

2. One of the endpoints of $\alpha_t$ may disappear in a death singularity.

In contrast, birth singularities do not pose a problem for the extension of the family of arcs past the time at which they occur.

Let $t_1$ be the time of the first death singularity. To avoid the two problems above we pick a small $\varepsilon > 0$ and at time $t_1' = t_1 - \varepsilon$ we jump from $\alpha_{t_1}^- := \alpha_{t_1}$ to a different arc $\alpha_{t_1}^+$ that also joins points of $g_{t_1}^{-1}(c)$ in distinct dotted components but that avoids a neighborhood of the death singularity. We will show how to construct $\alpha_{t_1}^+$ so that this jump can be filled in appropriately for the construction of the disk $E$. We will then extend the family of arcs $\alpha_t$ for $t > t_1'$ by starting with $\alpha_{t_1}^+$ and continuing past $t_1$ until just before the next death singularity occurs at some time $t_2 > t_1$. In the first case we show that the arc $\alpha_t$ can be replaced with one that avoids a neighborhood of the death singularity. In the second case we show that $\alpha_t$ can be replaced with an arc having an endpoint that avoids the death singularity.

For a continuous map $g : D \to \mathbb{R}^3$, we say that two arcs in the disk $D$ joining points of $g^{-1}(c)$ in distinct dotted components of $g^{-1}(T)$ are $g$-equivalent if their
images under $g$ are homotopic through arcs in $\mathbb{R}^3$ whose interiors are disjoint from $c$ and whose endpoints lie on $c$. Our goal is then to construct the arc $\alpha_{i_1}^+ \leftrightarrow$ so that it is $g_{i_1}$-equivalent to $\alpha_{i_1}^-$. When $g_t$ is transverse to $\partial T$ the dotted components of $g_{i_1}^{-1}(T)$ form planar subsurfaces of $D$, each a disk with holes. A boundary curve of a dotted component that has dots on both of its sides in $D$ is called primary, and boundary components with all dots on one side are called secondary.

If $\alpha_{i_1}^-$ leaves a dotted component $X$ through a secondary boundary component $b$, it must reenter $X$ through $b$, since $b$ is separating in $D$. Let $\beta$ be a subarc of $\alpha_{i_1}^-$ running between two successive intersections of $\alpha_{i_1}^-$ with $b$ and with interior outside of $X$. Then $\beta$ runs through either a dotless disc or a dotless annulus in $D$. We can then homotope $\beta$ into $b$ rel endpoints without crossing any dots. Push $\beta$ a little further into the interior of $X$. Repeating this process we can homotope $\alpha_{i_1}^-$ rel endpoints without crossing dots and so that it crosses only primary boundary components. See Figure 4. We abuse notation somewhat and continue to refer to this arc as $\alpha_{i_1}^-$. Our next goal is to arrange for $\alpha_{i_1}^-$ to pass through each primary boundary component exactly once. We will achieve this with the following lemma.

**Lemma 3.2.** Suppose $T$ is an embedded solid torus in $\mathbb{R}^3$ with core $c$, and suppose $g : D \to \mathbb{R}^3$ is a continuous map of a disc into $\mathbb{R}^3$ for which $g^{-1}(T)$ has two dotted components, each with image having algebraic intersection number $\pm 1$ with $c$. Let $\alpha$ be an arc in $D$, joining dots in distinct dotted components of $g^{-1}(T)$, and with

![Figure 4. Removing intersections of $\alpha_{i_1}^-$ with secondary boundary components of $g_{i_1}^{-1}(T)$](image-url)
Let $\beta \subset g^{-1}(T)$ be a subarc of $\alpha$ that starts and ends on the same primary boundary component $b$ of $g^{-1}(T)$. Then there is an arc $\beta' \subset b$ with the same endpoints as $\beta$ and with the property that replacing $\beta$ with $\beta'$ in $\alpha$ yields an arc $\alpha'$ that is $g$-equivalent to $\alpha$.

**Proof.** There is a homotopy of $\beta$ in $D$ rel endpoints to an arc $\bar{\beta}$ contained in $b$, possibly crossing dots. So $g(\beta)$ is homotopic rel endpoints in $\mathbb{R}^3$ to an arc $g(\bar{\beta})$ in $\partial T \cap g(D)$. This homotopy may pass outside $T$, as $\beta$ slides over holes of the dotted component. However the boundaries of these holes are secondary, since there are precisely two dotted components, and therefore bound disks in $\mathbb{R}^3$ disjoint from $c$. It follows that they have image under $g$ that is homotopically trivial on $\partial T$ or they have images on $\partial T$ that are nontrivial and bound disks in the exterior of $T$. In the latter case $T$ is unknotted, and we are done. Thus we can assume that $g(\beta)$ and $g(\bar{\beta})$ are homotopic rel endpoints in $T$. The arc $g(\beta)$ is also homotopic rel endpoints in $T - c$, by radial projection away from $c$, to an arc $\nu$ on $\partial T$. See Figure 5, which for clarity shows only part of $g(D)$.

Now, $g(\bar{\beta})$ and $\nu$ are homotopic rel endpoints in $T$ and so in $\partial T$ they differ by a multiple of a meridian. Note that the curve $g(b)$ is a meridian, since it bounds a disk in $T$ meeting $c$ algebraically once. So $g(\beta)$ can be homotoped rel endpoints in the complement of $c$ in $T$ to $\nu$, and then in turn to a curve formed by concatenating $g(\bar{\beta})$ with a multiple of $g(b)$. Take $\beta'$ to be $\bar{\beta}$ followed by this multiple of $b$. □

Now suppose that $\beta$ is a subarc of $\alpha_{\overline{i}}$ that enters and leaves a dotted component. Using Lemma 3.2 we can replace it with a $g_{\overline{i}}$-equivalent arc $\beta'$ that lies entirely on $g_{\overline{i}}^{-1}(\partial T)$, and then perturb $\beta'$ slightly so that it is disjoint from $g_{\overline{i}}^{-1}(T)$, as illustrated in Figure 6. In this way we replace $\alpha_{\overline{i}}$ with a $g_{\overline{i}}$-equivalent arc that has fewer intersections with dotted components, and by repeating we may remove all subarcs of $\alpha_{\overline{i}}$ that enter and leave a dotted component. We continue to refer to the resulting arc as $\alpha_{\overline{i}}$. Note that $\alpha_{\overline{i}}$ may now intersect itself.

We have found an arc $g_{\overline{i}}$-equivalent to the original arc $\alpha_{\overline{i}}$ that starts at a dot in one dotted component, exits that dotted component, then enters the second dotted component and finally ends at a dot. The following lemma allows us to
TOPOLOGICAL AND PHYSICAL LINK THEORY ARE DISTINCT

find a $g_{t_1^-}$-equivalent arc that replaces a subarc running from a primary boundary component to a dot, with any other arc running from the same point to a dot and not leaving $g_{t_1^-}(T)$.

We define an arc in a solid torus $T$ with core $c$ to be core-to-boundary if it has one endpoint on $\partial T$, the other endpoint on $c$, and interior disjoint from $c$.

Lemma 3.3. Let $T$ be a solid torus with core $c$. Let $\gamma$ and $\gamma'$ be core-to-boundary arcs in $T$ with $\gamma \cap \partial T = \gamma' \cap \partial T$. Then $\gamma$ and $\gamma'$ can be joined by a homotopy of core-to-boundary arcs, keeping the endpoint on $\partial T$ fixed.

Proof. We can lift $\gamma$ and $\gamma'$ to the universal cover of $T$, which is homeomorphic to $(\text{disk}) \times \mathbb{R}$, so that their common endpoint on $\partial T$ lifts to the same point $x$ while $c$ lifts to $\{0\} \times \mathbb{R}$. Each lift can be homotoped, keeping $x$ fixed and moving points only along the $\mathbb{R}$-factor, to the slice $(\text{disk}) \times \{\text{point}\}$ containing $x$. Further, the resulting arcs are homotopic rel endpoints in $(\text{disk}) \times \{\text{point}\}$ via arcs that miss $\{0\} \times \{\text{point}\}$ in their interior. Therefore the lifts of $\gamma$ and $\gamma'$ are homotopic through arcs joining $x$ to $\{0\} \times \mathbb{R}$ and with interiors disjoint from $\{0\} \times \mathbb{R}$. The projection of this homotopy to $T$ gives a homotopy joining $\gamma$ and $\gamma'$ through core-to-boundary arcs in $T$. \hfill \Box

Now suppose that a death singularity takes place in a dotted component containing a segment of $\alpha_{t_1^-}$. Let $\gamma_1^-'$ be the segment of $\alpha_{t_1^-}$, running from a dot to the primary boundary component. Choose an arc $\gamma_1'^-$ in the same dotted component that runs from a dot to the point where $\gamma_1^-$ exits the dotted component, so that $\gamma_1'^-$ is disjoint from a neighborhood containing the two dying dots. This is possible because the number of dots in each dotted component is odd. By Lemma 3.3, $\alpha_{t_1^-}$ is $g_{t_1^-}$-equivalent to the arc formed by replacing $\gamma_1^-$ by $\gamma_1'^-$. We take $\alpha_{t_1^-}$ to be the arc, $g_{t_1^-}$-equivalent to $\alpha_{t_1^-}$, which is obtained from $\alpha_{t_1^-}$ after making all these changes. See Figure 7. See also Figure 8 for an example, illustrated in $\mathbb{R}^3$, of how Lemma 3.3 may be applied.

We have constructed a family of arcs $\alpha_t$ in $D$ that varies continuously until time $t_1' = t_1 - \varepsilon$. It then jumps from $\alpha_{t_1^-}$ to the $g_{t_1^-}$-equivalent arc $\alpha_{t_1^+}^-$. The arc $\alpha_{t_1^+}$ was

![Figure 6](image-url)  
**Figure 6.** Removing an arc in $D$ that starts and ends on the same primary boundary component of $g^{-1}(T)$.  

![Figure 7](image-url)  
**Figure 7.** Illustrating the application of Lemma 3.3.

![Figure 8](image-url)  
**Figure 8.** Example of how Lemma 3.3 may be applied in $\mathbb{R}^3$. 


chosen to avoid a neighborhood of the death singularity at time $t_1$, so we can extend the family of arcs $\alpha_t$ past time $t = t_1$ until just before the next death singularity at time $t = t_2$. We then repeat this process.

Eventually, at time $t = t_l \in (0, 1)$ the two dotted components must merge along $g_l^{-1}(\partial T)$. At this time $g_l^{-1}(\partial T)$ consists of a collection of finitely many simple closed curves and one bouquet of finitely many circles embedded in $D$. The arc $\alpha_t$ begins and ends at a dot, and may run in and out of the single dotted component.

We now look at the complementary components in $D$ of the single dotted component of $g_l^{-1}(T)$. Each of these is either a disk or an annulus with $\partial D$ as one boundary component. Any subarc $\beta$ of $\alpha_t$ that runs out of a dotted component into a complementary component $X$ eventually leaves $X$ and reenters the dotted component. Since $X$ is either a disk or an annulus having $\partial D$ as one of its boundary
components, we can homotope $\beta$ rel endpoints off $X$ and into the dotted component without passing through any dots, since all dots lie within the single dotted component. In this way we can homotope $\alpha_t$, so that it lies entirely within the dotted component of $g_t^{-1}(T)$. This means that $g_t(\alpha_t)$ now lies entirely within $T$. It is then straightforward to shrink $g_t(\alpha_t)$ within $T$, keeping its interior disjoint from $c$ and its endpoints on $c$, until it collapses to a point on $c$.

We now form a spanning disk $E$ for $c$. Let $\beta_t = g_t \circ \alpha_t, \ t \in [0, t_1]$. Then $\beta_t$ is a family of arcs in $\mathbb{R}^3$ whose endpoints lie on $c$ and whose interiors are disjoint from $c$. These arcs vary continuously except at finitely many times $t'_1, t'_2, \ldots, t'_n$ just before death singularities. At these times the limiting arcs $\alpha^{-t'}_i$ and $\alpha^{+t'}_i$, as $t$ approaches $t'_i$ from below and above, are $g_{t'_i}$-equivalent. Finally, $\alpha_t$ is homotopic to a point on $c$ via arcs that start and end on $c$ but have interiors disjoint from $c$. Therefore there is a family of arcs, with endpoints on $c$ sweeping out a disk with interior in the complement of $c$, that represents a homotopy of $\beta_0$ to a point in $c$.

Let $\partial E$ denote the union of these arcs in $\mathbb{R}^3$ and take $E$ to be the disk obtained by taking the union of $\partial E$ with its reflection in the $xz$-plane. Note that the interior of $E$ does not intersect $c$ and that $\partial E \subseteq c$.

Let $a$ be one of the two points of intersection of $c$ with $g_0(D)$. Then $\partial E$ intersects $a$ in an odd number of points, one coming from $\beta_0$ and an additional even number coming from equal numbers of intersections of $a$ with $\partial E$ and its reflection. So $\partial E$ is nontrivial in $\pi_1(c)$. By the Loop Theorem [Papakyriakopoulos 1957], $c$ is the boundary of an embedded disc and therefore unknotted.

4. Proof of Theorem 1.2

We now prove Theorem 1.2, showing that if $L$ can be split via a physical isotopy then the length of $L_1$ must be at least $4\pi + 6$. Note that in this isotopy both components may move.

Assume there is a physical isotopy $I_s, s \in [0, 1], of \mathbb{R}^3$ with $I_0$ the identity, $I_1$ taking $L_1$ and $L_2$ to opposite sides of a plane, and with the length of the unknotted component $L_1$ being less than $4\pi + 6$. We will derive a contradiction.

During the course of the isotopy $I_s$ it is possible that the radius-one solid torus neighborhoods of the two link components bump against themselves or each other. We describe a slight modification of the isotopy that keeps the two components embedded and disjoint. Throughout the isotopy, tubular neighborhoods of any radius $r < 1$ give embedded disjoint solid torus neighborhoods of each of $L_1$ and $L_2$. Take $r = 1 - \epsilon'$ to be slightly less than 1 and then rescale the entire isotopy $I_s$ by $1/(1 - \epsilon')$. This restores the radius to 1 at the cost of slightly lengthening $L_1$ and $L_2$. With $\epsilon'$ small, the length of $L_1$ remains below $4\pi + 6$. The rescaled physical isotopy is then $(1 + \epsilon)$-thick, with $\epsilon = \epsilon'/(1 - \epsilon').
For a curve $c$ in $\mathbb{R}^3$, let $T(c)$ denote the radius-1 tubular neighborhood of $c$. Without loss of generality we can assume that the isotopy $I_s$ preserves $T(L_1)$ and $T(L_2)$, so that $I_5(T(L_i)) = T(I_5(L_i))$ for $i = 1, 2$. We then let $T_s$ denote the solid torus $T(I_s(L_2))$. For each $s \in [0, 1]$, let $x_s$ be the center of mass of the embedded curve $I_s(L_1)$ and let $f_s : D \to \mathbb{R}^3$ parametrize the disk forming the cone over $I_s(L_1)$ with cone point $x_s$. Since the cone point is inside the convex hull, its cone angle is always at least $2\pi$ [Cantarella et al. 2002; Gage 1980; Gromov 1983]. The maps $f_s : D \to \mathbb{R}^3$ induce a family of metrics on the disk $D$, parametrized by $s$, in which each disc is flat except at the cone point and is therefore a subdisk of a complete CAT(0) surface obtained by extending the rays from the cone point to infinity.

Now perturb $f_s$, $s \in [0, 1]$, so that the family of maps $f_s$ is generic, but leaving $f_s$ unchanged for $s$ in a small neighborhood of 0 and unchanged on $\partial D$ for all $s$. By generic, we mean that

1. $f_s$ is transverse to $c$ except for a finite number of times $s$ at which a birth or death of a pair of points of $f_s^{-1}(c)$ occurs;
2. $f_s$ is transverse to $\partial T$ at these times; and
3. $f_s$ is transverse to $\partial T$ except for a finite number of times when $f_s^{-1}(\partial T)$ consists of finitely many simple closed curves and a single component that is a bouquet of finitely many circles (in the case of a saddle-type singularity) or a single point (in the case of a birth or death singularity).

Genericity can be achieved by approximating the appropriate parts of $f_s$ by PL maps and using general position. The perturbation can be made arbitrarily $C^0$-small, and we let $f'_s$ denote the result of perturbing $f_s$ in this manner.

Each component of $f'_s^{-1}(T_s)$ in $D$ that contains a point of $f'_s^{-1}(I_s(L_2))$ contains a disc of radius 1 in $D$ enclosing that point, measured in the induced metric. The distance in $D$ of $\partial D$ from each of these components is at least 1. Suppose for a contradiction that there are three components of $f'_s^{-1}(T_s)$ containing a point of $f'_s^{-1}(I_s(L_2))$ for some $s$ and furthermore suppose that this is true no matter how small we made the perturbation of $f_s$ that gave $f'_s$. Then $f'_s^{-1}(T_s)$ contains three disks of radius 1 with disjoint interiors, and with $\partial D$ having distance at least 1 from each disk. This contradicts Proposition 2.1, since $L_1$ has length less than $4\pi + 6$. Hence, by taking the perturbation to obtain $f'_s$ to be sufficiently small, we can arrange that for all $s$ there are at most two components of $f'_s^{-1}(T_s)$ containing a point of $f'_s^{-1}(I_s(L_2))$.

We now define a family of disks $h_s : D \to \mathbb{R}^3$, $0 \leq s \leq 1$, by setting $h_s = I_s^{-1} \circ f'_s$. Extend $h_s$ to $-1 \leq s \leq 0$ by reflecting through the $xz$-plane, setting $h_s = r \circ h_{-s}$ for $s < 0$. Each $h_s$ maps $\partial D$ to $L_1$ and for $s \in [-1, 1]$ the preimage $h_s^{-1}(T(L_2))$ has at most two components containing a point of $h_s^{-1}(L_2)$. Moreover $h_0(D)$ lies
in the \(xz\)-plane, and \(h_1(D)\) and \(h_{-1}(D)\) are disjoint from \(L_2\). The disks \(h_3(D)\) now satisfy all the conditions of Proposition 3.1, implying that \(L_2\) is unknotted. This contradiction proves Theorem 1.2.

\[\Box\]

5. Some open problems

The following related problems remain open.

(1) Does there exist a Gordian unknot?

(2) Can the methods of Theorem 1.2 be extended to produce a Gordian split link with two unknotted components?

(3) In Theorem 1.2 the length of each component is fixed. One can formulate a problem where the sum of the component lengths is fixed but the individual components are allowed to stretch. Is there a Gordian split link in that setting?

Acknowledgements

We thank the anonymous referee for helpful comments that led to the improvement of this paper.

References


Received August 6, 2012. Revised September 10, 2013.

ALEXANDER COWARD
DEPARTMENT OF MATHEMATICS
UC BERKELEY
970 EVANS HALL #3840
BERKELEY, CA 94720
UNITED STATES
coward@math.berkeley.edu

JOEL HASS
MATHEMATICS DEPARTMENT
UNIVERSITY OF CALIFORNIA
DAVIS, CA 95616
UNITED STATES
hass@math.ucdavis.edu
Free evolution on algebras with two states, II
MICHAEL ANSHELEVICH

Systems of parameters and holonomicity of A-hypergeometric systems
CHRISTINE BERKESCH ZAMAERE, STEPHEN GRIFFETH and EZRA MILLER

Complex interpolation and twisted twisted Hilbert spaces
FÉLIX CABELLO SÁNCHEZ, JESÚS M. F. CASTILLO and NIGEL J. KALTON

The ramification group filtrations of certain function field extensions
JEFFREY A. CASTAÑEDA and QINGQUAN WU

A mean field type flow, II: Existence and convergence
JEAN-BAPTISTE CASTÉRAS

Isometric embedding of negatively curved complete surfaces in Lorentz–Minkowski space
BING-LONG CHEN and LE YIN

The complex Monge–Ampère equation on some compact Hermitian manifolds
JIANCHUN CHU

Topological and physical link theory are distinct
ALEXANDER COWARD and JOEL HASS

The measures of asymmetry for coproducts of convex bodies
QI GUO, JINFENG GUO and XUNLI SU

Regularity and analyticity of solutions in a direction for elliptic equations
YONGYANG JIN, DONGSHENG LI and XU-JIA WANG

On the density theorem for the subdifferential of convex functions on Hadamard spaces
MINA MOVAHEDI, DARYOUSH BEHMARDI and SEYEDEHSMAYEH HOSSEINI

$L^p$ regularity of weighted Szegő projections on the unit disc
SAMANGI MUNASINGHE and YUNUS E. ZEYTUNCU

Topology of complete Finsler manifolds admitting convex functions
SORIN V. SABAU and KATSUHIRO SHIOHAMA

Variations of the telescope conjecture and Bousfield lattices for localized categories of spectra
F. LUKE WOLCOTT