The homotopy and cohomology of spaces of locally convex curves in the sphere — II

Nicolau C. Saldanha

March 15, 2012

Abstract

A smooth curve $\gamma : [0,1] \to \mathbb{S}^2$ is locally convex if its geodesic curvature is positive at every point. J. A. Little showed that the space of all locally positive curves γ with $\gamma(0) = \gamma(1) = e_1$ and $\gamma'(0) = \gamma'(1) = e_2$ has three connected components $\mathcal{L}_{-1,c}$, $\mathcal{L}_{+1}$, $\mathcal{L}_{-1,n}$. The space $\mathcal{L}_{-1,c}$ is known to be contractible but the topology of the other two connected components is not well understood. We prove that all connected components of $\mathcal{L}_I$ are simply connected, that $H^2(\mathcal{L}_{+1};\mathbb{Z}) = \mathbb{Z}^2$ and $H^2(\mathcal{L}_{-1,n};\mathbb{Z}) = \mathbb{Z}$.

1 Introduction

A curve $\gamma : [0,1] \to \mathbb{S}^2$ is called *locally convex* if its geodesic curvature is always positive, or, equivalently, if $\det(\gamma(t), \gamma'(t), \gamma''(t)) > 0$ for all t. Let $\mathcal{L}_I$ be the space of all locally convex curves γ with $\gamma(0) = \gamma(1) = e_1$ and $\gamma'(0) = \gamma'(1) = e_2$. J. A. Little [2] showed that $\mathcal{L}_I$ has three connected components $\mathcal{L}_{-1,c}$, $\mathcal{L}_{+1}$, $\mathcal{L}_{-1,n}$: we call these the Little spaces. Figure 1 shows examples of curves in $\mathcal{L}_{-1,c}$, $\mathcal{L}_{+1}$ and $\mathcal{L}_{-1,n}$, respectively. The space $\mathcal{L}_{-1,c}$ is known to be contractible ([7]) but the topology of the other two connected components is not well understood. In this series of papers we present new results concerning the homotopy and cohomology of the Little spaces. A more ambitious aim would be to determine the homotopy type of these spaces (which we hope to accomplish in [3]).

Let $\mathcal{I}_I$ be the space of immersed curves $\gamma : [0,1] \to \mathbb{S}^2$, $\gamma(0) = \gamma(1) = e_1$, $\gamma'(0) = \gamma'(1) = e_2$. For each $\gamma \in \mathcal{I}_I$, consider its Frenet frame $\mathfrak{F}_\gamma : [0,1] \to SO(3)$

2000 *Mathematics Subject Classification.* Primary 57N65, 53C42; Secondary 34B05. *Keywords and phrases* Convex curves, topology in infinite dimension, periodic solutions of linear ODEs.

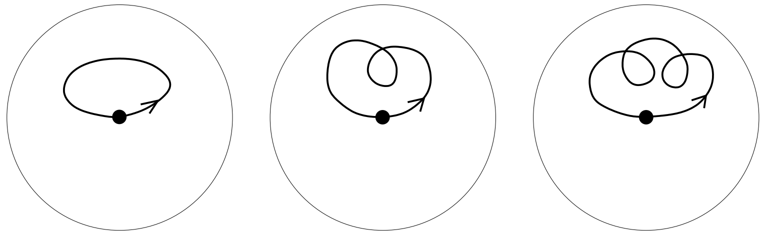


Figure 1: Curves in $\mathcal{L}_{-1,c}$, $\mathcal{L}_{+1}$ and $\mathcal{L}_{-1,n}$.

and its lift $\tilde{\mathfrak{F}}_\gamma : [0,1] \to \mathbb{S}^3$. The value of $\tilde{\mathfrak{F}}_\gamma(1)$ defines the two connected components of $\mathcal{I}_I$: $\gamma \in \mathcal{I}_{+1}$ if and only if $\tilde{\mathfrak{F}}_\gamma(1) = 1$. It is well know that each space $\mathcal{I}_{\pm 1}$ is homotopically equivalent to $\Omega\mathbb{S}^3$. In particular each $\mathcal{I}_{\pm 1}$ is connected, simply connected and there is an element $\mathbf{x} \in H^2(\mathcal{I}_{\pm 1};\mathbb{Z})$ such that each $H^{2k}(\mathcal{I}_{\pm 1};\mathbb{Z}) = \mathbb{Z}$ is generated by $\mathbf{x}^k$. Let $\mathcal{L}_{+1} = \mathcal{I}_{+1} \cap \mathcal{L}_I$, $\mathcal{L}_{-1} = \mathcal{L}_{-1,c} \sqcup \mathcal{L}_{-1,n} = \mathcal{I}_{-1} \cap \mathcal{L}_I$. In the first paper ([4]) we saw that the inclusions $\mathcal{L}_{\pm 1} \subset \mathcal{I}_{\pm 1}$ are homotopically surjective but not homotopy equivalences. Indeed, we constructed elements $\mathbf{f}_{2k} \in H^{2k}(\mathcal{L}_{(-1)^{(k+1)}};\mathbb{Z})$ and maps $\mathbf{g}_{2k} : \mathbb{S}^{2k} \to \mathcal{L}_{(-1)^{(k+1)}}$ with $\mathbf{f}_{2k}(\mathbf{g}_{2k}) = 1$, $\mathbf{g}_{2k}$ homotopic to a constant in $\mathcal{I}_{(-1)^{(k+1)}}$ and $\mathbf{f}_{2k}$ not in (the image of) $H^{2k}(\mathcal{I}_{(-1)^{(k+1)}};\mathbb{Z})$. In other words, there we give lower estimates for the groups $H^{2k}(\mathcal{L}_{\pm 1};\mathbb{Z})$ and $\pi_{2k}(\mathcal{L}_{\pm 1})$. In the present paper we give upper estimates which imply the following theorem.

Theorem 1 *The connected components of $\mathcal{L}_I$ are simply connected. Furthermore, $H^2(\mathcal{L}_{-1,n};\mathbb{Z})$ is generated by $\mathbf{x}$ and $H^2(\mathcal{L}_{+1};\mathbb{Z})$ is generated by $\mathbf{x}$ and $\mathbf{f}_2$. Also, $\pi_2(\mathcal{L}_{-1,n}) = \mathbb{Z}$ and $\pi_2(\mathcal{L}_{+1}) = \mathbb{Z}^2$ is generated by $\mathbf{g}_2$ and $\tilde{\mathbf{g}}_2$.*

Notice that $\mathcal{L}_{per}$, the set of all 1-periodic locally convex curves $\tilde{\gamma} : \mathbb{R} \to \mathbb{S}^2$ is homeomorphic to $SO(3) \times \mathcal{L}_I$: define $\Psi : \mathcal{L}_{per} \to SO(3) \times \mathcal{L}_I$ by $\Psi(\tilde{\gamma}) = (\mathfrak{F}_{\tilde{\gamma}}(0), (\mathfrak{F}_{\tilde{\gamma}}(0))^{-1}\tilde{\gamma}|_{[0,1]})$. We usually prefer to work in $\mathcal{L}_I$ but sometimes move to $\mathcal{L}_{per}$.

In Section 2 we review some known results. Section 3 contains an algebraic description of the all-important construction $\Delta^\sharp$; pulling one loop around is a special case of $\Delta^\sharp$. Section 4 discusses the uses and limitations of $\Delta^\sharp$ to prove that a map $f : K \to \mathcal{L}_I$ is homotopic to $\nu_2 * f$ (which essentially reduces the problem to the well-understood scenario of immersions). In Section 5 the discussion becomes more geometric and less algebraic as we discuss loops and the set $\mathcal{T}_0 \subset \mathcal{L}_{+1}$ of stars: roughly, once $\mathcal{T}_0$ is removed, $\mathcal{L}_{+1} \smallsetminus \mathcal{T}_0$ rather resembles $\mathcal{I}_{+1}$. Section 6 polishes a few nasty configurations so that we can complete the proof of our main results in Section 7. Section 8 is a very short conclusion.

The author would like to thank Dan Burghelea and Boris Shapiro for helpful conversations. The author acknowledges the hospitality of The Mathematics

Department of The Ohio State University during the winter quarters of 2004 and 2009 and the support of CNPq, Capes and Faperj (Brazil).

2 Previous results

One of the fundamental constructions in Little's argument is that if the curve $\gamma \in \mathcal{L}_I$ has a loop, we can add a pair of loops as in Figure 2: in (a), the loop moves one full turn along a geodesic and in (b) the large loops are shrunk (we will discuss (c) later).

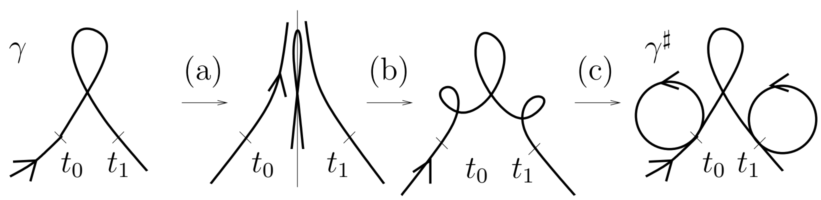


Figure 2: How to go from γ to $\gamma^\sharp$

Let $\mathcal{C}_0$ be the circle with diameter e_1e_3, parametrized by $\nu_1 \in \mathcal{L}_I$,

$$\nu_1(t) = \left(\frac{1+\cos(2\pi t)}{2}, \frac{\sqrt{2}}{2}\sin(2\pi t), \frac{1-\cos(2\pi t)}{2} \right).$$

For positive n, let $\nu_n(t) = \nu_1(nt)$ so that $\nu_1 \in \mathcal{L}_{-1,c}$ and, for $n > 1$, $\nu_n \in \mathcal{L}_{(-1)^n}$.

For $\gamma_1 \in \mathcal{I}_{\sigma_1}$, $\gamma_2 \in \mathcal{I}_{\sigma_2}$, $\sigma_i \in \{+1, -1\}$, let $\gamma_1 * \gamma_2 \in \mathcal{I}_{\sigma_1\sigma_2}$ be defined by

$$(\gamma_1 * \gamma_2)(t) = \begin{cases} \gamma_1(2t), & 0 \le t \le 1/2, \\ \gamma_2(2t-1), & 1/2 \le t \le 1. \end{cases}$$

Notice that if $\gamma_1, \gamma_2 \in \mathcal{L}_I$ then $\gamma_1 * \gamma_2 \in \mathcal{L}_I$. For $f : K \to \mathcal{I}_I$, let $\nu_n * f : K \to \mathcal{I}_I$ be defined by $(\nu_n * f)(p) = \nu_n * (f(p))$. Intuitively, $\nu_n * f$ is obtained from f by adding n loops to $f(p)$ at the point $f(p)(0)$. We may want to spread out n loops along the curve: for $\gamma \in \mathcal{I}_\sigma$ and for large n, define $(F_n(\gamma))(t) = \mathfrak{F}_\gamma(t)\nu_n(t)$. For small n, the above function from $[0,1]$ to $\mathbb{S}^2$ may not be an immersion. For sufficiently large n, however, $F_n(\gamma) \in \mathcal{L}_{(-1)^n\sigma}$.

By the above construction, it is clear that given $\gamma \in \mathcal{L}_{+1} \sqcup \mathcal{L}_{-1,n}$ there exist $H_a, H_b : [0,1] \to \mathcal{L}_{\pm 1}$, $H_a(0) = H_b(0) = \gamma$, $H_a(1) = \nu_2 * \gamma$, $H_b(1) = F_{2n}(\gamma)$. The construction is not uniform, however: is depends on the choice of the loop. In other words, given a compact set K and a map $f : K \to \mathcal{L}_{+1} \sqcup \mathcal{L}_{-1,n}$, the existence of $H : [0,1] \times K \to \mathcal{L}_{+1} \sqcup \mathcal{L}_{-1,n}$, $H(0,\cdot) = f$, $H(1,\cdot) = \nu_2 * f$ or $H(1,\cdot) = F_{2n} \circ f$ is not clear at this point. Indeed, the existence (or not) of such a homotopy is the crucial point in this paper. The following proposition helps clarify the situation.

Proposition 2.1 ([4]) *Let K be a compact set and let $f : K \to \mathcal{L}_I \subset \mathcal{I}_I$ a continuous function.*

(a) For sufficiently large n, the functions $\nu_2 * f$ and $F_{2n} \circ f$ are homotopic in $\mathcal{L}_I$.

(b) If $f = \nu_1 * \tilde{f}$ (for some $\tilde{f}$) then f is homotopic to $\nu_2 * f$.

(c) The function f is homotopic to a constant in $\mathcal{I}_I$ if and only if $\nu_2 * f$ is homotopic to a constant in $\mathcal{L}_I$.

(d) There exists a map $\mathbf{g}_2 : \mathbb{S}^2 \to \mathcal{L}_{+1}$ such that $\nu_2 * \mathbf{g}_2$ is homotopic to a constant in $\mathcal{L}_{+1}$ but $\mathbf{g}_2$ is not.

3 Bruhat cells and the set $\mathcal{W}$

Given a locally convex curve $\gamma : [0,1] \to \mathbb{S}^2$ and a 3×3 matrix A with positive determinant, the curve $\gamma^A : [0,1] \to \mathbb{S}^2$,

$$\gamma^A(t) = \frac{A\gamma(t)}{|A\gamma(t)|}$$

is also locally convex. Furthermore, $\mathfrak{F}_{\gamma^A}(t) = A\mathfrak{F}_\gamma(t)U$ for $U \in \mathcal{U}^+$, where $\mathcal{U}^+$ is the group of upper triangular 3×3 matrices with positive off-diagonal entries.

Let $\mathcal{U}^1 \subset \mathcal{U}^+$ be the group of upper triangular matrices with unit diagonal. Recall that $SO(3)$ is divided in *Bruhat cells* by the following equivalence relation: Q_1 and Q_2 are equivalent if and only if there exist $U_1 \in \mathcal{U}^1$ and $U_2 \in \mathcal{U}^+$ with $Q_1 = U_1 Q_2 U_2$. The group Weyl group $D_3 \subset SO(3)$ of signed permutation matrices with positive determinant has one element per cell. The four open cells $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4$ have respective representatives

$$J_1 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$J_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Further recall ([6], [7], [5]) that given $Q \in SO(3)$ there exists a convex curve $\gamma : [0,1]$ with $\mathfrak{F}_\gamma(0) = I$ and $\mathfrak{F}_\gamma(1) = Q$ if and only if Q belongs to $\mathcal{J}_2$ or to one of the 5 lower dimensional cells in its boundary corresponding to the following matrices:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

As a consequence, given $Q \in SO(3)$, there exists $\gamma \in \mathcal{L}_{-1,c}$ with $\mathfrak{F}_\gamma(1/2) = Q$ if and only if $Q \in \mathcal{J}_2$. Similarly, $Q_0^{-1}Q_1 \in \mathcal{J}_4$ if and only if there exists a convex curve $\gamma : [0,1] \to \mathbb{S}^2$ with $\gamma(0) = \gamma(1) = Q_0 e_1$, $\gamma'(0) = \gamma'(1) = -Q_0 e_2$, $\gamma(1/2) = Q_1 e_1$, $\gamma'(1/2) = -Q_1 e_2$.

Let $\mathcal{W} \subset \mathcal{L}_I \times \mathbb{S}^1 \times \mathbb{S}^1$ be the set of triples (γ, t_0, t_1) such that $(\mathfrak{F}_\gamma(t_0))^{-1}\mathfrak{F}_\gamma(t_1) \in \mathcal{J}_4$. Define $U_1 : \mathcal{W} \to \mathcal{U}^1$ and $U_2 : \mathcal{W} \to \mathcal{U}^+$ so that $(\mathfrak{F}_\gamma(t_0))^{-1}\mathfrak{F}_\gamma(t_1) = U_1 J_4 U_2$. Alternatively, $(\gamma, t_0, t_1) \in \mathcal{W}$ if and only if there exists a convex curve $\alpha : \mathbb{S}^1 \to \mathbb{S}^2$, $\alpha(0) = \gamma(t_0)$, $\alpha'(0) = -\gamma'(t_0)$, $\alpha(1) = \gamma(t_1)$, $\alpha'(1) = -\gamma'(t_1)$. Notice that $\mathcal{W}$ is an open subset of $\mathcal{L}_I \times \mathbb{S}^1 \times \mathbb{S}^1$; $(\gamma, t_0, t_1) \in \mathcal{W}$ implies $t_0 \neq t_1$ and $(\gamma, t_1, t_0) \in \mathcal{W}$.

We now define the function $\Delta^\sharp : [0, +\infty) \times \mathcal{W} \to \mathcal{L}_I$, one of our main technical tools throughout the paper.

Definition 3.1 *Let $(\gamma, t_0, t_1) \in \mathcal{W}$ and $t_\bullet \in \mathbb{S}^1$ with $t_\bullet < t_0 < t_1 < t_\bullet + 1$. Take $\epsilon^\sharp > 0$, $\epsilon^\sharp < (1/20)\min(t_0 - t_\bullet, t_1 - t_0, t_\bullet + 1 - t_1)$. Let $Q_0 = \mathfrak{F}_\gamma(t_0)$, $Q_1 = \mathfrak{F}_\gamma(t_1)$, $U_1 = U_1(\gamma, t_0, t_1)$, $U_2 = U_2(\gamma, t_0, t_1)$ so that $Q_0^{-1}Q_1 = U_1 J_4 U_2$. Let $\alpha = \gamma^{U_1^{-1}Q_0^{-1}}$, i.e.,*

$$\alpha(t) = \frac{U_1^{-1}Q_0^{-1}\gamma(t)}{|U_1^{-1}Q_0^{-1}\gamma(t)|};$$

α is locally convex with $\mathfrak{F}_\alpha(t_0) = I$, $\mathfrak{F}_\alpha(t_1) = J_4$. For $s \in [0, +\infty)$, let $T_s = \min(s, 2)$; define an increasing piecewise linear homeomorphisms $h_{a,s} : [0, T_s\epsilon^\sharp] \to [0, s]$ whose graph is a polygonal line with vertices $(0,0)$, $(\epsilon^\sharp, 1)$ (if $s \geq 1$) and $(T_s\epsilon^\sharp, s)$. Similarly, let the graph of $h_{b,s} : [t_0 + \epsilon^\sharp T_s, t_1 - \epsilon^\sharp T_s] \to [t_0, t_1]$ have vertices $(t_0 + T_s\epsilon^\sharp, t_0)$, $(t_0 + 4\epsilon^\sharp, t_0 + 4\epsilon^\sharp)$ $(t_1 - 4\epsilon^\sharp, t_1 - 4\epsilon^\sharp)$ and $(t_1 - T_s\epsilon^\sharp, t_1)$. Define $\alpha_s : [t_0, t_1] \to \mathbb{S}^2$ by

$$\alpha_s(t) = \begin{cases} \nu_1\left(h_{a,s}(t - t_0)\right), & t_0 \leq t \leq t_0 + T_s\epsilon^\sharp, \\ \mathfrak{F}_{\nu_1}(s)\alpha\left(h_{b,s}(t)\right), & t_0 + T_s\epsilon^\sharp \leq t \leq t_1 - T_s\epsilon^\sharp, \\ J_4\nu_1\left(-h_{a,s}(t_1 - t)\right), & t_1 - T_s\epsilon^\sharp \leq t \leq t_1. \end{cases}$$

Finally,

$$Q\Delta^\sharp(s, \gamma, t_0, t_1) = \begin{cases} \gamma(t), & t_\bullet \leq t \leq t_0 \text{ or } t_1 \leq t \leq t_\bullet, \\ (\alpha_s)^{Q_0 U_1}(t), & t_0 \leq t \leq t_1, \end{cases}$$

where $Q \in SO(3)$ is uniquely chosen so that $\mathfrak{F}_{\Delta^\sharp(s,\gamma,t_0,t_1)}(0) = I$.

A few remarks are in order. The curve α_s is obtained from α by attaching an arc of circle of angle $2\pi s$ to either end, rotating the curve to keep the same endpoints and reparametrizing. Similarly, $\Delta^\sharp(s, \gamma, t_0, t_1)$ is obtained from γ by inserting an arc of $2\pi s$ at positions t_0 and t_1; if s is not an integer, the portion of γ between t_0 and t_1 will be "rotated". Up to minor deformations, the path $\Delta^\sharp(s, \gamma, t_0, t_1)$, $s \in [0,1]$, from γ to $\gamma^\sharp_{t_0,t_1} = \Delta^\sharp(1, \gamma, t_0, t_1)$ is exemplified in Figure

2 (in a situation where $0 < t_0 < t_1 < 1$). The loop between t_0 and t_1 is pushed along a geodesic all the way, until it comes back (a). The two long chunks of curve (in the figure, very nearly geodesics) are then shrunk (b) and rounded (c) so that we obtain $\gamma^\sharp_{t_0,t_1}$. Notice that the portion of γ outside the interval $[t_0, t_1]$ is unaltered throughout the process.

Strictly speaking, the definition of $\Delta^\sharp$ depends on the choice of $\epsilon^\sharp$: the only difference, however, when you change $\epsilon^\sharp$ is that functions get reparametrized. Similarly, we ask that $U_1 \in \mathcal{U}^1$ so that U_1 and U_2 become uniquely determined. The function h_s in Definition 3.1 is chosen so that the following technical result holds.

Lemma 3.2 *Let K be a compact set and $(f, t_0, t_1) : K \to \mathcal{W}$ a continuous map; there exists $\epsilon^\sharp > 0$ which suits the definition of $\Delta^\sharp(s, f(p), t_0(p), t_1(p))$ for all s.*

Let $\mathcal{K} \subset \mathcal{L}_I$ be the compact set of all curves of the form $\Delta^\sharp(s, f(p), t_0(p), t_1(p))$, $p \in K$, $s \in [0, 3]$. For $s > 0$, let $n = \lfloor s - 3 \rfloor$, $\tilde{s} = s - n$. For $p \in K$, let $\gamma = \Delta^\sharp(s, f(p), t_0(p), t_1(p))$, $\tilde{\gamma} = \Delta^\sharp(\tilde{s}, f(p), t_0(p), t_1(p))$. Let

$$h(t) = \begin{cases} 0, & t_\bullet \le t \le t_0, \\ n(t - t_0)/\epsilon^\sharp, & t_0 \le t \le t_0 + \epsilon^\sharp, \\ n, & t_0 + \epsilon^\sharp \le t \le t_1 - \epsilon^\sharp, \\ 2n - n(t_1 - t)/\epsilon^\sharp, & t_1 - \epsilon^\sharp \le t \le t_1, \\ 2n, & t_1 \le t \le t_\bullet. \end{cases}$$

Then $\tilde{\gamma} \in \mathcal{K}$ and

$$\gamma(t) = Q\mathfrak{F}_{\tilde{\gamma}}(t)\nu_1^{U_1(p)}(h(t)), \quad Q \in SO(3),$$

where $U_1(p) = U_1(f(p), t_0(p), t_1(p))$.

Proof: This is a straightforward computation. ■

The functions $\Delta^\sharp(\tilde{s}, (\Delta^\sharp(s, \gamma, t_0, t_1), t_0, t_1))$ and $\Delta^\sharp(s + \tilde{s}, \gamma, t_0, t_1)$ differ by reparametrization only. Under suitable hypothesis, a related identity holds for distinct points (γ, t_0, t_1), (γ, t_2, t_3). Two points $(\gamma, t_0, t_1), (\gamma, t_2, t_3) \in \mathcal{W}$ are *disjoint* if the intervals $[t_0, t_1], [t_2, t_3] \subset \mathbb{S}^1$ are disjoint, or, equivalently, if $t_0 < t_1 < t_2 < t_3 < t_0 + 1$ or $t_2 < t_3 < t_0 < t_1 < t_2 + 1$.

Lemma 3.3 *If $(\gamma, t_0, t_1), (\gamma, t_2, t_3) \in \mathcal{W}$ are disjoint then, for any $s, \tilde{s} \in [0, 1]$,*

$$(\Delta^\sharp(s, \gamma, t_0, t_1), t_2, t_3), (\Delta^\sharp(\tilde{s}, \gamma, t_2, t_3), t_0, t_1) \in \mathcal{W}.$$

Furthermore,

$$\Delta^\sharp(\tilde{s}, (\Delta^\sharp(s, \gamma, t_0, t_1), t_2, t_3)) = \Delta^\sharp(s, (\Delta^\sharp(\tilde{s}, \gamma, t_2, t_3), t_0, t_1)).$$

Proof: This follows directly from the construction of $\Delta^\sharp$ and of the function γ^*_s in the definition. Indeed, γ^*_s for (t_0, t_1) coincides with γ in an open interval containing (t_2, t_3) (and vice versa). ■

4 Disjoint covers

Recall that one of our aims is to decide whether a continuous map $f : K \to \mathcal{L}_I$ is homotopic to $\nu_2 * f$. In this section, we present several situations where this is the case and one example where this is not the case. We start with a simple example.

Lemma 4.1 *Let K be a compact set and $f : K \to \mathcal{L}_I$ a continuous map. If there is a continuous function $t_1 : K \to \mathbb{S}^1$ such that for all $p \in K$ we have $(f(p), 0, t_1(p)) \in \mathcal{W}$ then f is homotopic to $\nu_2 * f$.*

Proof: Let $H : [0,1] \times K \to \mathcal{L}_I$ be defined by $H(s,p) = \Delta^\sharp(2s, f(p), 0, t_1(p))$. Up to reparametrization, $H(1,p) = \nu_2 * \tilde{f}(p)$, where $\tilde{f}(p)$ is obtained from $f(p)$ by inserting two turns at $t_1(p)$. Since ν_2 and ν_4 are in the same connected component, $\nu_2 * \tilde{f}$ is homotopic to $\nu_2 * (\nu_2 * \tilde{f})$ and therefore to $\nu_2 * f$. ■

Given this result, some questions are natural:

- If there exist continuous functions $t_0, t_1 : K \to \mathbb{S}^1$ with $(f(p), t_0(p), t_1(p)) \in \mathcal{W}$, does it follow that f is homotopic to $\nu_2 * f$?
- If for every p there exist $t_0, t_1 \in \mathbb{S}^1$ such that $(f(p), t_0, t_1) \in \mathcal{W}$, does it follow that f is homotopic to $\nu_2 * f$?

As we shall see, the answers are yes and no, respectively. Before we attack these problems, however, we introduce a few concepts.

Define
$$\mathcal{O} = \{\gamma \in \mathcal{L}_I \mid \forall t_0, t_1 \in \mathbb{S}^1, (\gamma, t_0, t_1) \notin \mathcal{W}\} :$$
the set $\mathcal{O}$ is clearly a closed subset of $\mathcal{L}_I$. A *double point* of a curve $\gamma \in \mathcal{L}_I$ is a pair $(t_0, t_1) \in (\mathbb{S}^1)^2$, $t_0 \neq t_1$, with $\gamma(t_0) = \gamma(t_1)$. Similarly, an *$n$-tuple point* is an n-tuple $(t_0, t_1, \ldots, t_{n-1})$, $0 \leq t_0 < t_1 < \cdots < t_{n-1} < 1$, such that $\gamma(t_0) = \gamma(t_1) = \cdots = \gamma(t_{n-1})$. We identify the double point (t_1, t_0) with (t_0, t_1). A double point (t_0, t_1) is a *self-tangency* if $\gamma'(t_0)$ and $\gamma'(t_1)$ are parallel and *transversal* otherwise. A self-tangency (t_0, t_1) is *positive* if $\gamma'(t_1)$ is a positive multiple of $\gamma'(t_0)$ (and *negative* otherwise).

If (t_0, t_1) is a positive self-tangency, the normal vector $\mathbf{n}_\gamma(t) = \mathfrak{F}_\gamma(t) e_3$ satisfies $\mathbf{n}_\gamma(t_0) = \mathbf{n}_\gamma(t_1)$. Define $h_1(t) = \langle \gamma(t), \mathfrak{F}_\gamma(t) e_2 \rangle$, $h_2(t) = \langle \gamma(t), \mathbf{n}_\gamma(t_0) \rangle$: notice that $h_1(t_0) = h_1(t_1) = h_2(t_0) = h_2(t_1) = 0$, $h_2'(t_0) = h_2'(t_1) = 0$, $h_1'(t_0) > 0$ and $h_1'(t_1) > 0$. There exists $\epsilon > 0$ such that $h_1|_{(t_0-\epsilon, t_0+\epsilon)}$ and $h_1|_{(t_1-\epsilon, t_1+\epsilon)}$ are invertible. Let $g_0 = h_2 \circ (h_1|_{(t_0-\epsilon, t_0+\epsilon)})^{-1}$ and $g_1 = h_2 \circ (h_1|_{(t_1-\epsilon, t_1+\epsilon)})^{-1}$: near the origin, the graph of g_i is the orthogonal projection of the image of the curve γ near $\gamma(t_i)$ to the tangent plane, using $\gamma'(t_i)/|\gamma'(t_i)|$ and $\mathbf{n}_\gamma(t_i)$ as basis. By construction, $g_0(0) = g_1(0) = 0$ and $g_0'(0) = g_1'(0) = 0$. By convexity, $g_0''(0) > 0$,

$g_1''(0) > 0$. The self-tangency (t_0, t_1) has *order* n if $g_0^{(n)}(0) \neq g_1^{(n)}(0)$ but $g_0^{(j)}(0) = g_1^{(j)}(0)$ for any $j < n$ and order $+\infty$ if $g_0^{(j)}(0) = g_1^{(j)}(0)$ for any j. A *self-osculating* point is a positive self-tangency of order 3 or more.

Lemma 4.2 *Let (t_0, t_1) be a double point of $\gamma \in \mathcal{L}_I$. If (t_0, t_1) is either transversal or a negative self-tangency then for any $\epsilon > 0$ there exist $\tilde{t}_0, \tilde{t}_1$, $|\tilde{t}_0 - t_0| + |\tilde{t}_1 - t_1| < \epsilon$, $(\gamma, t_0, t_1) \in \mathcal{W}$.*

Proof: First consider transversal double points. Assume without loss of generality that $\det(\gamma(t_0), \gamma'(t_0), \gamma'(t_1)) > 0$. For small ϵ we may take $\tilde{t}_0 = t_0 + \epsilon/4$, $\tilde{t}_1 = t_1 - \epsilon/4$: this is a straightforward computation but is probably best verified geometrically in Figure 3: the dashed convex curves tangent to γ validate the geometric characterization of $\mathcal{W}$.

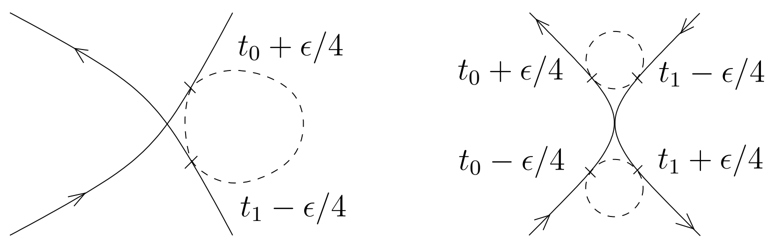


Figure 3: Obtaining $(\tilde{t}_0, \tilde{t}_1)$ such that $(\gamma, \tilde{t}_0, \tilde{t}_1) \in \mathcal{W}$.

For a negative self-tangency and small ϵ we may take either $\tilde{t}_0 = t_0 + \epsilon/4$, $\tilde{t}_1 = t_1 - \epsilon/4$ or $\tilde{t}_0 = t_0 - \epsilon/4$, $\tilde{t}_1 = t_1 + \epsilon/4$: this is again verified in Figure 3. ■

It follows directly from this result that if $\gamma \in \mathcal{O}$ then all double points of γ are positive self-tangencies.

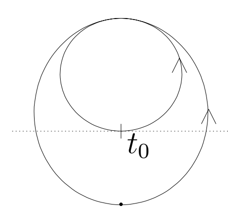


Figure 4: A curve in $\mathcal{O}_2$.

The following proposition settles the second question raised at the beginning of this section. Recall ([4]) that $\mathbf{g}_{+,2} : \mathbb{S}^2 \to \mathcal{L}_{+1}$ is an explicit function such that $\mathbf{g}_{+,2}$ and $\nu_2 * \mathbf{g}_{+,2}$ are not homotopic in $\mathcal{L}_{+1}$ (even though they are homotopic in $\mathcal{I}_{+1}$).

Proposition 4.3 *There is a map $f : \mathbb{S}^2 \to \mathcal{L}_{+1} \smallsetminus \mathcal{O}$ which is homotopic (in $\mathcal{L}_{+1}$) to $\mathbf{g}_2$. In particular, f and $\nu_2 * f$ are not homotopic.*

Proof: Let $\mathbb{S}^1$ be the unit circle in the complex plane and let $D \subset \mathbb{C}$ be the closed disk of radius $1/4$. For $s \in D$, let $g_s : \mathbb{S}^1 \to \mathbb{R}^2$ be defined by

$$g_s(z) = -iz^2 - \frac{i}{10}z^{-3} + sz^{-1};$$

some such curves are drawn in Figure 5 (the curve corresponding to $s = 0$ is in position $(6, 5)$, i.e., sixth row, fifth column).

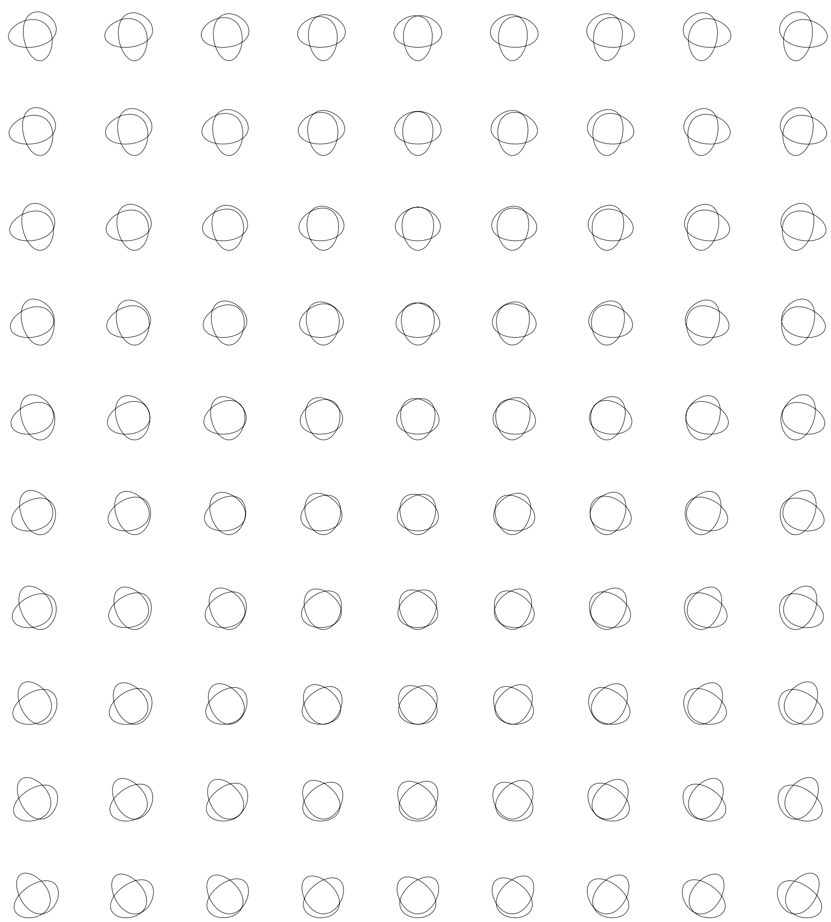


Figure 5: A family of convex curves.

A straightforward computation verifies that these curves are locally convex in the plane: central projection obtains a similar family of locally convex curves in the sphere. Notice that there are 5 curves with self-osculating points (approximately in positions $(3, 5)$, $(5, 2)$, $(5, 8)$, $(9, 3)$ and $(9, 7)$) but none of them lie in $\mathcal{O}$ since they all have transversal double points.

Remove a small disk near ν_2 in the function $\mathbf{g}_{+,2}$ corresponding to the region below the bottom row in Figure 9 of [4]; the above family can be used to plug the hole. The resulting function if homotopic to $\mathbf{g}_{+,2}$ and its image is contained in $\mathcal{L}_{+1} \smallsetminus \mathcal{O}$. ■

Before we answer the first question, we present another situation where f is guaranteed to be homotopic to $\nu_2 * f$.

Lemma 4.4 *Let K be a compact manifold and $f : K \to \mathcal{L}_I$ be a continuous map. Assume there exist functions $t_0, t_1 : K \to (0,1)$ such that, for all $p \in K$:*

(a) $0 < t_0(p) < t_1(p) < 1$*;*

(b) $f(p)|_{[0,t_0(p)]}$ *and* $f(p)|_{[t_0(p),t_1(p)]}$ *are convex;*

(c) there exists a convex curve $\alpha(p) : [0,1] \to \mathbb{S}^2$ *with* $\mathfrak{F}_{\alpha(p)}(0) = \mathfrak{F}_{f(p)}(t_0(p))$, $\mathfrak{F}_{\alpha(p)}(1/2) = \mathfrak{F}_{f(p)}(0)$, $\mathfrak{F}_{\alpha(p)}(1) = \mathfrak{F}_{f(p)}(t_1(p))$.

Then f is homotopic to $\nu_2 * f$.

Proof: Define $H : [0,1] \to \mathcal{L}_I$ for $s \in [0,1/2]$ with $H(0,p) = f(p)$ and

$$H(1/2,p)(t) = \begin{cases} f(p)(t), & t \in [0,t_0(p)] \cup [t_1(p),1], \\ \alpha(p)\left(\frac{t-t_0(p)}{t_1(p)-t_0(p)}\right), & t \in [t_0(p),t_1(p)]. \end{cases}$$

Contractibility of the space of convex curves with prescribed initial and final value and direction guarantees that this can be done. Let $t_{1/2}(0) = (t_0(p)+t_1(p))/2$: notice that $H(1/2,p)|_{[0,t_{1/2}(p)]}$ is a closed convex curve. Set

$$H(1,p)(t) = \begin{cases} \nu_1(2t), & t \in [0,1/2], \\ H(1/2,p)\left(\frac{t-t_{1/2}(p)}{1-t_{1/2}(p)}\right), & t \in [1/2,1]; \end{cases}$$

contractibility of $\mathcal{L}_{-1,c}$ guarantees that this can be done. Now $\tilde{f} : K \to \mathcal{L}_I$, $\tilde{f}(p) = H(1,p)$, is of the form $\tilde{f} = \nu_1 * (\text{something})$. The result now follows from Proposition 2.1, item (b). ■

For a function $f : K \to \mathcal{L}_I$, a finite open cover $K = \bigcup_{i=1,\ldots,N} V_i$ together with functions $t_{0,i}, t_{1,i} : V_i \to \mathbb{S}^1$ is a *disjoint cover of f* if:

(a) if $p \in V_i$ then $(f(p), t_{0,i}(p), t_{1,i}(p)) \in \mathcal{W}$;

(b) if $p \in V_i \cap V_j$ then the points $(f(p), t_{0,i}(p), t_{1,i}(p)), (f(p), t_{0,j}(p), t_{1,j}(p)) \in \mathcal{W}$ are either equal or disjoint.

The following lemma settles the fist question; it is actually much stronger.

Lemma 4.5 *Let K be a compact manifold and $f : K \to \mathcal{L}_I$ be a continuous map. If the function f admits a disjoint cover then f is homotopic to $\nu_2 * f$.*

It is instructive to verify directly that the function f constructed in Proposition 4.3 does not admit a disjoint cover.

Proof: Consider a function f and a cover by disjoint loops with the notation above. Recall that the support $\operatorname{supp}(\phi)$ of $\phi : K \to [0,1]$ is the closure of $\phi^{-1}((0,1])$. Let $\phi_i : K \to [0,1]$ be a partition of unit: $\operatorname{supp}(\phi_i) \subset V_i$ and $\sum_i \phi_i(p) = 1$.

Intuitively, our first step is to apply $\Delta^\sharp(\sigma, \cdots)$ to the points given by the cover; σ is a function which goes to zero together with ϕ_i. More precisely, select $\epsilon^\sharp > 0$ such that if $\phi_{i_1}(p), \phi_{i_2}(p) > 0$ and $t_{0,i_1}(p) < t_{1,i_1}(p) < t_{0,i_2}(p) < t_{1,i_2}(p) < 1 + t_{0,i_1}(p)$ then

$$20\epsilon^\sharp < \min(t_{1,i_1}(p) - t_{0,i_1}(p), t_{0,i_2}(p) - t_{1,i_1}(p), t_{1,i_2}(p) - t_{0,i_2}(p), 1 + t_{0,i_1}(p) - t_{1,i_2}(p)).$$

For $p \in K$, let $(t_{0,j,p}, t_{1,j,p})$, $j = 1, \ldots, N_p$, $1 \le N_p \le N$, be the distinct pairs for which there exists $i \in \{1, \ldots, N\}$, $p \in V_i$, $t_{0,j,p} = \tilde{t}_{0,i}(p)$, $t_{1,j,p} = \tilde{t}_{1,i}(p)$. For each j, let $I_j \subset \{1, \ldots, N\}$ be the set of indices i for which the above conditions hold. Define

$$\psi_j(p) = \sum_{i \in I_j} \phi_i(p).$$

For $s \in [0, 1/2]$, let $\sigma_j(s,p) = \min(2M, 12MNs\psi_j(p))$ where $M > 0$ is a large integer to be specified later. Let $H_0(s,p) = f(p)$ and define recursively

$$H_j(s,p) = \Delta^\sharp(\sigma_j(s,p), H_{j-1}(s,p), t_{0,j,p}, t_{1,j,p})$$

and $H(s,p) = H_{N_p}(s,p)$. By Lemma 3.3 the order of the indices does not matter.

Intuitively, we added many turns to each curve and must now spread them. Define $U_1 : K \times \mathbb{S}^1 \to \mathcal{U}^1$ so that if $p \in V_i$, $t \in [t_{0,i}(p), t_{0,i}(p) + 4\epsilon^\sharp] \cup [t_{1,i}(p) - 4\epsilon^\sharp, t_{1,i}(p)]$ then $U_1(p,t) = U_1(f(p), t_{0,i}(p), t_{1,i}(p))$. Write

$$R(p,t,x) = U_1(p,t)\mathfrak{F}_{\nu_1}(x)(U_1(p,t))^{-1};$$

notice that $R(p,t,x_1)R(p,t,x_2) = R(p,t,x_1+x_2)$. For any $p \in K$ there is at least one index j such that $\psi_j(p) \ge 1/(3N)$ and therefore $\sigma_i(1/2,p) = 2M$: assume without loss of generality that these indices are $j = 1, \ldots, \tilde{N}_p$, $1 \le \tilde{N}_p \le N_p$. We define an auxiliary curve $\eta(p) \in \mathcal{L}_I$ by $\eta_0 = f(p)$, $\eta_j = \Delta^\sharp(M, \eta_{j-1}, t_{0,j,p}, t_{1,j,p})$ for $j = 1, \ldots, \tilde{N}_p$ and $\eta_j = \Delta^\sharp(\sigma_j(1/2,p), \eta_{j-1}, t_{0,j,p}, t_{1,j,p})$ for $j = \tilde{N}_p + 1, \ldots, N_p$ and $\eta(p) = \eta_{N_p}$ so that, by Lemma 3.2,

$$H(1/2,p)(t) = r(p,t)\mathfrak{F}_{\eta(p)}(t) R\left(p, t, M \sum_{j=1,\ldots,\tilde{N}_p} (\beta_{0,j}(t) + \beta_{1,j}(t))\right) e_1,$$

where $r(p,t) \in (0,+\infty)$ and

$$\beta_{0,j} = \beta_{[t_{0,j,p},t_{0,j,p}+\epsilon^\sharp]}, \quad \beta_{1,j} = \beta_{[t_{1,j,p}-\epsilon^\sharp,t_{1,j,p}]}, \quad \beta_{[t_-,t_+]}(t) = \begin{cases} 0, & t \le t_-, \\ \frac{t-t_-}{t_+-t_-}, & t_- \le t \le t_+, \\ 1, & t \ge t_+. \end{cases}$$

Given an interval $[t_-,t_+] \subset \mathbb{S}^1$ define $t_\bullet = (t_+ + t_- - 1)/2$ so that $t_\bullet < t_- < t_+ < t_\bullet + 1$. Given $\theta \in [0,1]$, let

$$[t_-,t_+]^\theta = [\theta t_\bullet + (1-\theta)t_-, \theta(t_\bullet + 1) + (1-\theta)t_+]$$

so that if $\theta = 1$ the interval degenerates to the whole circle. Let $\theta_j : [1/2,1]\times K \to [0,1]$ be

$$\theta_j(s,p) = \begin{cases} 0, & \psi_j(p) \le \frac{1}{3N}, \\ (2s-1)(3N\psi_j(p)-1), & \frac{1}{3N} \le \psi_j(p) \le \frac{2}{3N}, \\ 2s-1, & \psi_j(p) \ge \frac{2}{3N}. \end{cases}$$

Define

$$\beta_{0,j,s} = \beta_{[t_{0,j,p},t_{0,j,p}+\epsilon^\sharp]^{\theta_j(s,p)}}, \quad \beta_{1,j,s} = \beta_{[t_{1,j,p}-\epsilon^\sharp,t_{1,j,p}]^{\theta_j(s,p)}},$$

$$\begin{aligned} H(s,p)(t) &= r(p,s,t)\mathfrak{F}_{\eta(p)}(t) R\left(p,t,M\sum_{j=1,\ldots,\tilde{N}_p}(\beta_{0,j,s}(t)+\beta_{1,j,s}(t))\right)e_1 \\ &= r(p,s,t)\mathfrak{F}_{\eta(p)}(t)U_1(p,t)\nu_1\left(M\sum_{j=1,\ldots,\tilde{N}_p}(\beta_{0,j,s}(t)+\beta_{1,j,s}(t))\right) \end{aligned}$$

where $r(p,s,t)$ is a positive number chosen so that the expression has absolute value 1. For sufficiently large M, all the functions constructed above will belong to $\mathcal{L}_I$, as required. Indeed, $\gamma = H(1/2,p)$ is of the form $\gamma(t) = \gamma_1(t)/|\gamma_1(t)|$, $\gamma_1(t) = \mathfrak{F}_\eta(t)U_1(p,t)\nu_1(h(t))$ where η belongs to a compact set $\mathcal{K} \subset \mathcal{L}_I$ independent of the choice of M. For $s \in [1/2,1]$, $\gamma = H(s,p)$ still has the same form; for any given t, s and p, either $h(t) \in \mathbb{Z}$, $h'(t) = 0$ or $h'(t) > M$ (up to a few transition points which need not concern us). In the first case, local convexity of γ follows from local convexity of $H(1/2,p)$. In the second case, expanding $\det(\gamma_1(t),\gamma_1'(t),\gamma_1''(t))$ shows that this expression is positive provided M is large enough.

Let $\tilde{f} : K \to \mathcal{L}_I$, $\tilde{f}(p) = H(1,p)$. We claim that if M is large enough then $\tilde{f}$ satisfies the hypothesis of Lemma 4.4. Indeed, for γ in the image of $\tilde{f}$, let

$$\begin{pmatrix} g_1(t) \\ g_2(t) \\ g_3(t) \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \\ -\sqrt{2}/2 & 0 & \sqrt{2}/2 \end{pmatrix} (U_1(p,0))^{-1}(\mathfrak{F}_\gamma(0))^{-1}\gamma(t).$$

From the above form for γ we have, for small t,

$$\begin{pmatrix} g_1(t) \\ g_2(t) \\ g_3(t) \end{pmatrix} \approx \frac{\sqrt{2}}{2} r(p,s,t) \begin{pmatrix} \cos(2\pi h(t)) \\ \sin(2\pi h(t)) \\ 1 \end{pmatrix}.$$

Let $t_0(p)$ be the smallest $t > 0$ for which $g_1(t) = 0$, $g_2(t) < 0$; let $t_1(p)$ be the smallest $t > t_0(p)$ for which $g_1(t) = 0$, $g_2(t) > 0$. For sufficiently large M these are continuously defined and the convexity hypothesis will hold. This completes the proof of the claim and of the lemma. ■

5 Loops and stars

From now on our aim is to produce disjoint covers for functions $f : K \to \mathcal{L}_I$ or, at least, to prove that f is homotopic to $\tilde{f}$ such that $\tilde{f}$ admits a disjoint cover. We must therefore turn to the geometry of curves.

We now define nested dense open sets $\mathcal{L}_{\pm 1}^{(k)} \subset \mathcal{L}_{\pm 1}$ for $k \leq 3$; the complement $\mathcal{L}_{\pm 1} \smallsetminus \mathcal{L}_{\pm 1}^{(k)}$ has codimension k.

Let $\mathcal{L}_{\pm 1}^{\langle 0 \rangle} = \mathcal{L}_{\pm 1}^{(1)} \subset \mathcal{L}_{\pm 1}$ be the set of curves with no triple points or self-tangencies. Notice that the sets $\mathcal{L}_{\pm 1}^{(1)}$ have infinitely many connected components since the number of double points does not change in a connected component of these sets. Let $\mathcal{L}_{\pm 1}^{\langle 1,a \rangle} \subset \mathcal{L}_{\pm 1}$ be the set of curves with exactly one self-tangency of order 2, no triple points and no self-osculating points (and an arbitrary number of double points). Let $\mathcal{L}_{\pm 1}^{\langle 1,b \rangle} \subset \mathcal{L}_{\pm 1}$ be the set of curves with exactly one triple point and no self-tangencies. The sets $\mathcal{L}_{\pm 1}^{\langle 1,a \rangle}, \mathcal{L}_{\pm 1}^{\langle 1,b \rangle} \subset \mathcal{L}_{\pm 1}$ are disjoint submanifolds of codimension 1. Generically, the passage from one connected component of $\mathcal{L}_{\pm 1}^{(1)}$ to another crosses $\mathcal{L}_{\pm 1}^{\langle 1,* \rangle}$ transversally and is a Reidemeister move of type II (resp. III) if $* = a$ (resp. $* = b$; [1]); Reidemeister moves of type I are not allowed in $\mathcal{L}_I$. Figure 6 shows the possible Reidemeister moves in $\mathcal{L}_I$.

Define $\mathcal{L}_{\pm 1}^{(2)} = \mathcal{L}_{\pm 1}^{(1)} \sqcup \mathcal{L}_{\pm 1}^{\langle 1 \rangle} \subset \mathcal{L}_{\pm 1}$. Let $\mathcal{L}_{\pm 1}^{\langle 2,* \rangle}$ be the set of curves having:

(a) exactly one self-tangency, which is positive and of order 3, and no triple points;

(b) exactly two self-tangencies, both of order 2, and no triple points;

(c) exactly one triple point where there is also a self-tangency of order 2;

(d) exactly two (unrelated) triple points;

(e) exactly one quadruple point;

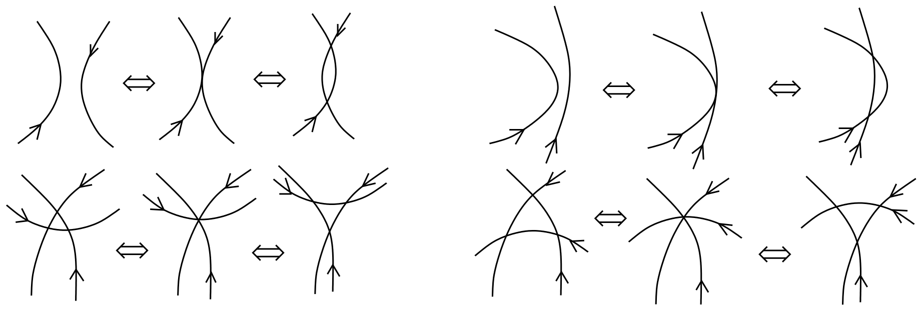


Figure 6: Reidemeister moves: type II on first line, type III on second line.

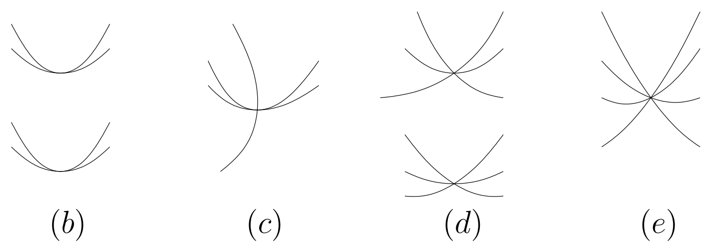


Figure 7: Curves in $\mathcal{L}_{\pm1}^{\langle 2,*\rangle}$.

Figure 7 illustrates these situations. These sets are submanifolds of codimension 2. Finally, define $\mathcal{L}_{\pm1}^{(3)} = \mathcal{L}_{\pm1}^{(2)} \sqcup \mathcal{L}_{\pm1}^{\langle 2,a\rangle} \sqcup \cdots \sqcup \mathcal{L}_{\pm1}^{\langle 2,e\rangle} \subset \mathcal{L}_{\pm1}$.

We already saw examples curves in $\mathcal{L}_{+1}^{\langle 2,a\rangle}$ in Figures 4 and 5; the second one shows a surface transversal to $\mathcal{L}_{+1}^{\langle 2,a\rangle}$. Notice that the self-osculating point may be perturbed to become one single transversal double point or three transversal double points (observe the central column); the reader should compare this with perturbations of the real polynomial $P(x) = x^3$, which may admit one or three real roots.

A *loop* of a curve $\gamma \in \mathcal{L}_I$ is a transversal double point (t_0, t_1) such that the restriction $\gamma|_{[t_0,t_1)}$ is injective. We sometimes think of the loop as the interval $[t_0, t_1]$, the restriction of γ to this interval or even the image of this restriction. A loop is *direct* (resp. *reverse*) if $\det(\gamma(t_0), \gamma'(t_0), \gamma'(t_1))$ is negative (resp. positive). Figure 8 shows examples of loops.

A *star* is a curve γ in the same connected component of $\mathcal{L}_{+1}^{(1)}$ as one of the infinite family of curves given in Figure 9. More precisely, a star has $2k+1$ double points; if $k > 0$, their images in the sphere are the vertices of a convex polygon and, for any pair of adjacent vertices, there are two arcs of γ joining them. Alternatively, a star is a curve in $\mathcal{L}_{+1}^{(1)}$ which admits loops (t_0, t_1) and $(t_1, t_0 + 1)$.

Let $\mathcal{T}_0$ be the closure (in $\mathcal{L}_{+1}$) of the set of stars and let $\mathcal{T}_1$ be its boundary.

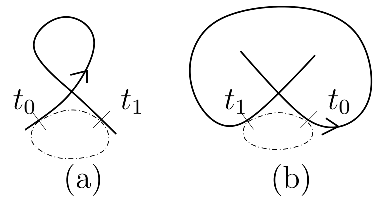


Figure 8: A direct loop and a reverse loop.

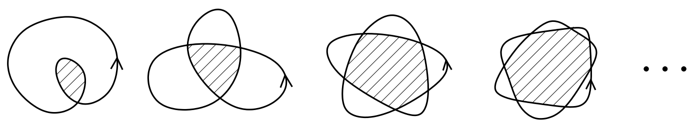


Figure 9: Stars ($k = 0, 1, 2, 3, \ldots$).

A curve $\gamma \in \mathcal{L}_{+1}^{\langle 1,b \rangle}$ with triple point (t_0, t_1, t_2) is a *trefoil* if (t_0, t_1), (t_1, t_2) and $(t_2, t_0 + 1)$ are direct loops.

Lemma 5.1 *The set $\mathcal{T}_1$ is the set of trefoils and is a manifold of codimension 1. The set $\mathcal{T}_0$ is contractible and $\mathcal{T}_1$ is homotopically equivalent to $\mathbb{S}^1$.*

Proof: We have to show that the only Reidemeister moves from a star to a generic γ which is not a star pass through a trefoil. In order to do this, we classify all possible Reidemeister moves starting at a star. Figure 10 shows how a Reidemeister move of type II takes a star to another star (changing the value of k) and how a Reidemeister move of type III takes a star ($k = 1$) to a generic curve which is not a star passing through a trefoil. We prove that these are the only possible moves.

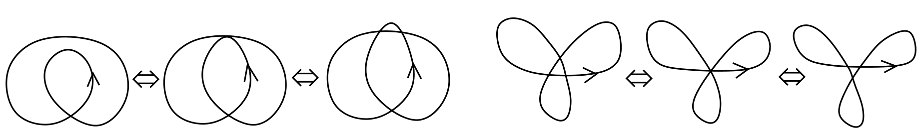


Figure 10: Reidemeister moves starting at a star.

The only possible star from which a Reidemeister move of type III is possible is the one shown in figure 10 ($k = 1$): indeed, a Reidemeister move of type III is quite impossible if the curve does not form a combinatorial triangle. In order to see that the only possible Reidemeister moves of type II are those indicated in figure 10, notice that if γ is a star, its image is trapped in the union of triangles shown in figure 11 (where straight lines indicate geodesics in the sphere).

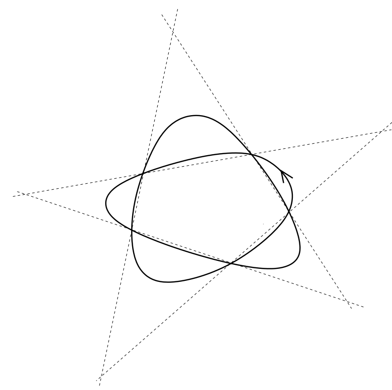


Figure 11: A star is trapped in a union of triangles.

For $\gamma \in \mathcal{L}_{+1}^{(1)} \cap \mathcal{T}_0$, let $K(\gamma) \subset \mathbb{S}^2$ be the closure of the region positively surrounded by γ. The set $K(\gamma)$ is shaded in Figure 9. Clearly, $K(\gamma)$ is a convex set. The definition of $K(\gamma)$ can be continuously extended to $\mathcal{T}_0$: for $\gamma \in \mathcal{T}_1$, $K(\gamma)$ consists of the triple point only. Let $k(\gamma) \in K(\gamma)$ be the baricenter of $K(\gamma)$ (recall that in order to find the baricenter of a convex subset of $\mathbb{S}^2$ we first find its baricenter in $\mathbb{R}^3$ and then radially project it onto the sphere). For $\gamma \in \mathcal{T}_1$, $k(\gamma)$ is the triple point. Figure 11 shows that $\gamma(t) \neq \pm k(\gamma)$ for any $\gamma \in \operatorname{int}(\mathcal{T}_0)$ and any t.

We construct a homotopy $H : [0,1] \times \operatorname{int}(\mathcal{T}_0) \to \operatorname{int}(\mathcal{T}_0)$, $H(0,\gamma) = \gamma$, $H(1,\cdot)$ constant equal to ν_2. Given γ, let v_1, v_2, v_3 be the only positively oriented orthonormal basis with $v_3 = k(\gamma)$, $\gamma(0)$ in the plane spanned by v_1 and v_3. Reparametrizing, we may write

$$\gamma(t) = \frac{1}{\sqrt{1 + (u(t))^2}} \, \left((\cos(4\pi t))v_1 + (\sin(4\pi t))v_2 + u(t)v_3 \right).$$

The condition for such a curve to be locally convex is that $u''(t) + 16\pi^2 u(t) > 0$. Set $u(s,t) = su(t) + (1-s)$ and (up to base point)

$$H(s,\gamma)(t) = \frac{1}{\sqrt{1 + (u(s,t))^2}} \, \left((\cos(4\pi t))v_1 + (\sin(4\pi t))v_2 + u(s,t)v_3 \right).$$

By the linearity of the above condition, all such curves are locally convex.

We construct the universal cover of $\mathcal{T}_1$. Let $\tilde{\mathcal{T}}_1$ be the set of pairs $(\tilde{\gamma}, \tilde{t}_0)$ where $\tilde{\gamma} : \mathbb{R} \to \mathbb{S}^2$ is a 1-periodic locally convex function with $\mathfrak{F}_{\tilde{\gamma}}(0) = I$, the restriction $\gamma = \tilde{\gamma}|_{[0,1]}$ belongs to $\mathcal{T}_1$ and $\tilde{t}_0 \in \mathbb{R}$ is a triple point, i.e., there exist $\tilde{t}_1, \tilde{t}_2 \in RR$, $\tilde{t}_0 < \tilde{t}_1 < \tilde{t}_2 < \tilde{t}_0 + 1$, $\tilde{\gamma}(\tilde{t}_0) = \tilde{\gamma}(\tilde{t}_1) = \tilde{\gamma}(\tilde{t}_2)$. The projection $\Pi : \tilde{\mathcal{T}}_1 \to \mathcal{T}_1$ takes $(\tilde{\gamma}, \tilde{t}_0) \in \tilde{\mathcal{T}}_1$ to $\gamma = \tilde{\gamma}|_{[0,1]} \in \mathcal{T}_1$. This is a covering map by construction; the group of deck transformations is isomorphic to $\mathbb{Z}$, spanned by $(\tilde{\gamma}, \tilde{t}_0) \mapsto (\tilde{\gamma}, \tilde{t}_1)$ (where $\tilde{t}_1$ is defined as above).

We claim that $\tilde{\mathcal{T}}_1$ is contractible. A homotopy $H : [0,1] \times \tilde{\mathcal{T}}_1 \to \tilde{\mathcal{T}}_1$ taking $\tilde{\mathcal{T}}_1$ to a point starts with, for $s \in [0,1/4]$,

$$H(s, \tilde{\gamma}, \tilde{t}_0) = (\tilde{\gamma}_s, (1-4s)\tilde{t}_0), \quad \tilde{\gamma}_s(t) = (\mathfrak{F}_{\tilde{\gamma}(4s\tilde{t}_0)})^{-1}\tilde{\gamma}(t + 4s\tilde{t}_0).$$

This defines a deformation retract from $\tilde{\mathcal{T}}_1$ to $\mathcal{F}_2$, the set of flowers with 3 petals ([4]). We now use the interval $s \in [1/4, 1/2]$ to reparametrize our curves so that for $s = 1/2$ the triple point of $H(s, \tilde{\gamma}, \tilde{t}_0)$ will be $t_0 = 0$, $t_1 = 1/3$, $t_2 = 2/3$. Next, for $s \in [1/2, 3/4]$, set $H(s, \tilde{\gamma}, \tilde{t}_0) = (H(1/2, \tilde{\gamma}, \tilde{t}_0))^U$, $U = U(\tilde{\gamma}, \tilde{t}_0) \in \mathcal{U}^1$, so that

$$\mathfrak{F}_{H(3/4,\tilde{\gamma},\tilde{t}_0)}(1/3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathfrak{F}_{H(3/4,\tilde{\gamma},\tilde{t}_0)}(2/3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

Finally, the loops $H(3/4, \tilde{\gamma}, \tilde{t}_0)|_I$, $I = [0,1/3], [1/3,2/3], [2/3,1]$, are convex: we use $s \in [3/4, 1]$ to deform them to some fixed loop. This completes the proof of the claim. Thus $\pi_1(\mathcal{T}_1) = \mathbb{Z}$ and $\mathcal{T}_1$ has a contractible universal cover, proving that $\mathcal{T}_1$ is homotopically equivalent to $\mathbb{S}^1$ and completing the proof of the lemma. ■

6 Eggs

A curve $\gamma \in \mathcal{L}_I$ is an *odd egg* if there exists a reverse loop (t_0, t_1) such that the image of γ is contained in the closed disk positively surrounded by the loop (Figure 12, (a)). Notice that a star with a single double point is an odd egg. A curve $\gamma \in \mathcal{L}_I$ is an *even egg* if there exist two transversal double points (t_0, t_1) and (t_2, t_3) with $t_0 < t_1 < t_2 < t_3 < t_0 + 1$, $\det(\gamma(t_0), \gamma'(t_0), \gamma'(t_1)) < 0$ and $\det(\gamma(t_2), \gamma'(t_2), \gamma'(t_3)) < 0$ such that $\gamma|_{[t_1,t_2)\cup[t_3,t_0+1)}$ is injective and the image of γ is contained in the closed disk positively surrounded by the above restriction (Figure 12, (b)). Let $\mathcal{E} \subset \mathcal{L}_I$ be the set of all eggs (even and odd).

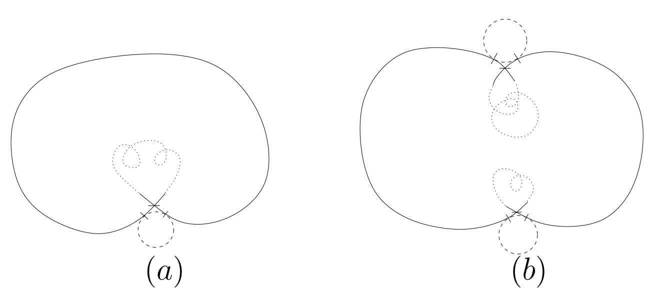


Figure 12: Odd and even eggs.

Lemma 6.1 *Let K be a compact manifold and $f : K \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$ a continuous map. Then f is homotopic in $\mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$ to some map $\hat{f} : K \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$.*

Proof: Intuitively, we pull the creature out of the egg.

There exist open sets $B_1 \subset B_2 \subset K$, $\overline{B_1} \subset B_2$, such that:

(a) if $p \notin B_1$ then $f(p)$ is not an odd egg;

(b) there exist functions $t_0, t_1 : B_2 \to \mathbb{S}^1$ with $(f(p), t_0(p), t_1(p)) \in \mathcal{W}$ for all $p \in B_2$;

(c) if $f(p)$ is an odd egg then $f(p)(t_0(p))$ and $f(p)(t_1(p))$ are approximately equal to the extrema of the reverse loop in the definition of odd eggs;

The functions t_0 and t_1 are indicated in Figure 12(a). Let $\phi : K \to [0,1]$ be a continuous function with $\phi|_{B_1} = 1$, $\phi|_{K \smallsetminus B_2} = 0$. Define

$$H(s,p) = \begin{cases} f(p), & p \notin B_2, \\ \Delta^\sharp(s\phi(p), f(p), t_0(p), t_1(p)), & p \in B_2. \end{cases}$$

This opens all odd eggs.

In order to get rid of the even eggs the construction is similar but with a harmless subtlety. We can easily define $B_3 \subset B_4 \subset K$, $\overline{B_3} \subset B_4$, such that if $p \notin B_3$ then $f(p)$ is not an even egg. In B_4 is not simply connected, it is not clear, however, that continuous functions $t_0, t_1, t_2, t_3 : B_4 \to \mathbb{S}^1$ can be defined since the two double points in the shell of the egg can trade places. Define therefore $t_0, t_1, t_2, t_3 : \tilde{B}_4 \to \mathbb{S}^1$, where $\tilde{B}_4$ is an appropriate double cover of B_4. Finally, define

$$H(s,p) = \begin{cases} f(p), & p \notin B_4, \\ \Delta^\sharp(s\phi(p), \Delta^\sharp(s\phi(p), f(p), t_0(\tilde{p}), t_1(\tilde{p})), t_2(\tilde{p}), t_3(\tilde{p})), & p \in B_4, \end{cases}$$

where $\tilde{p} \in \tilde{B}_4$ is one of the two lifts of p; Lemma 3.3 guarantees that both choices of $\tilde{p}$ obtain the same value for H. ∎

Lemma 6.2 *If $\gamma \in \mathcal{L}_I^{(3)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ then γ has at least one direct loop; if $\gamma \in \mathcal{L}_I^{(2)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ then γ has at least two direct loops.*

Proof: We first consider the case $\gamma \in \mathcal{L}_I^{(1)} \smallsetminus \mathcal{L}_{-1,c}$. Take $t_* \in \mathbb{S}^1$. Let

$$t_1 = \sup\{t \in \mathbb{S}^1 \mid \gamma|_{[t_*,t]} \text{ is injective}\}.$$

There exists a unique $t_0 \in [t_*, t_1)$ with $\gamma(t_0) = \gamma(t_1)$. The desired loop is (t_0, t_1). Still in $\mathcal{L}_I^{(1)} \smallsetminus \mathcal{L}_{-1,c}$, we prove the existence of a direct loop. Let (t_c, t_b) be a reverse loop. As in Figure 13 (a), draw a geodesic tangent to γ at $\gamma(t_b)$: the geodesic transversally intersects the image of γ at $\gamma(t_a)$, $t_a \in (t_c, t_b)$. Take $t_* \in (t_c, t_a)$: we claim that the construction above obtains a direct loop. More generally, assume the restriction of γ to $[t_*, t_b]$ is as in Figure 13 (b): an injective function such that the geodesic tangent to the image of γ at $\gamma(t_b)$ meets the image of γ transversally at $\gamma(t_a)$, $t_a \in [t_*, t_b)$; also, the image under γ of $[t_a, t_b]$ plus the segment of geodesic between $\gamma(t_b)$ and $\gamma(t_a)$ form the boundary of a convex closed disk $D(t_b) \subset \mathbb{S}^2$. Then, as $\tilde{t}_b$ increases starting from t_b the above condition in preserved (and $D(\tilde{t}_b)$ becomes smaller) until $\gamma(\tilde{t}_a) = \gamma(\tilde{t}_b)$, obtaining a direct loop.

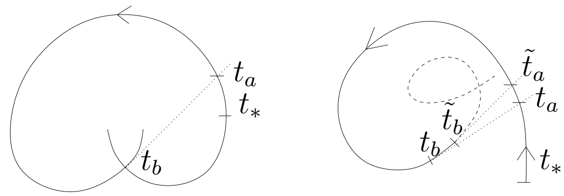


Figure 13: A configuration which obtains a direct loop.

The same construction and argument may be applied with time reversed: it follows that the only curves in $\mathcal{L}_I^{(1)} \smallsetminus \mathcal{L}_{-1,c}$ with a unique direct loop are those for which both constructions (original and with reversed time) lead to the same direct loop. Thus, the only curves with a unique direct loop are odd eggs.

If γ has no self-tangencies of odd order, perturb it near each self-tangency so as to destroy the self-tangency without creating new self-intersections. Consider a loop of the modified curve $\hat{\gamma}$: we claim that the same double point is a loop for the original curve γ. It suffices to show that self-tangencies can not be created within loops. For direct loops this follows from the convexity of the restriction. For a reverse loop (t_0, t_1), take the two geodesics tangent to γ at t_0 and t_1 as in Figure 14 and a point $\gamma(t_2)$ between the intersections of these geodesics with the image of γ. Consider a projective transformation taking the tangent geodesic to γ at $\gamma(t_2)$ to infinity and the geodesic joining $\gamma(t_0) = \gamma(t_1)$ with $\gamma(t_2)$. The image of γ under this projective transformation is the graph of a function, completing the proof in this case.

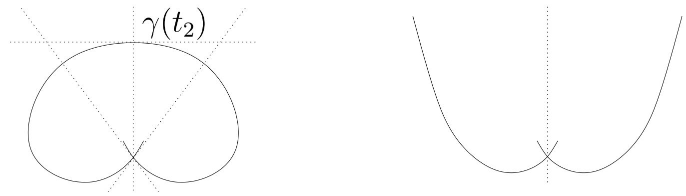


Figure 14: A reverse loop and its image under a projective transformation.

If γ has a single self-tangency (t_0, t_1) of order 3 start by perturbing a neighborhood of the self-tangency in such a way as to create exactly one transversal double point. Any direct loop of $\hat{\gamma}$ except for (t_0, t_1) (if it is simple) yields a direct loop of γ. Thus, the only situation where γ does not have a direct loop is if (t_0, t_1) is the unique direct loop of $\hat{\gamma}$; this completes the proof. ■

Two loops (t_0, t_1) and $(\tilde{t}_0, \tilde{t}_1)$ are *disjoint* if the intervals $[t_0, t_1] \subset \mathbb{S}^1$ and $[\tilde{t}_0, \tilde{t}_1] \subset \mathbb{S}^1$ are disjoint. Notice that this does not mean that the images of the intervals under γ are disjoint.

A curve $\gamma \in \mathcal{L}_I^{(1)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ is a *pseudo-egg* if it belongs to the connected component of one of the curves in the infinite family indicated in Figure 15. More precisely, γ admits two non disjoint direct loops (t_0, t_1) and (t_2, t_3), $t_0 < t_2 < t_1 < t_3 < t_0 + 1$, and the restriction of γ to $[t_3, t_0 + 1]$ is injective. Let $\mathcal{E}^+$ be the set of eggs and pseudo-eggs.

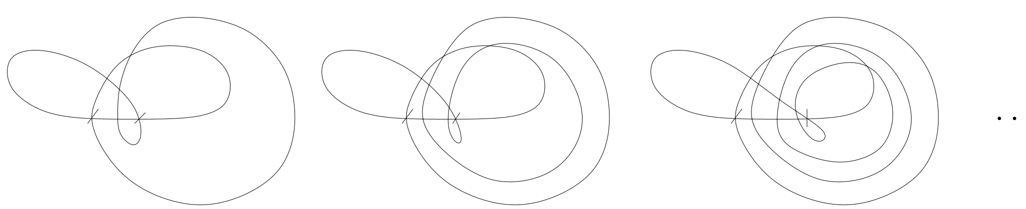


Figure 15: Pseudo-eggs

Lemma 6.3 *Let K be a compact manifold and $f : K \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ a continuous map. Then f is homotopic in $\mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ to some map $\hat{f} : K \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E}^+)$.*

Proof: Consider the innermost loop in the spiral and pull it (using $\Delta^\sharp$) as in Figure 16: this will either do a Reidemeister move of type III or of type II, in either case destroying the pseudo-egg. ■

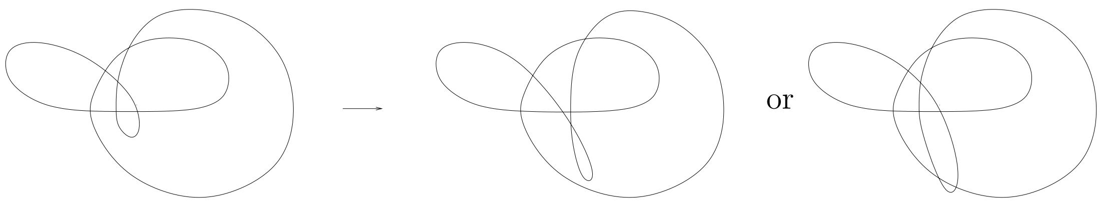
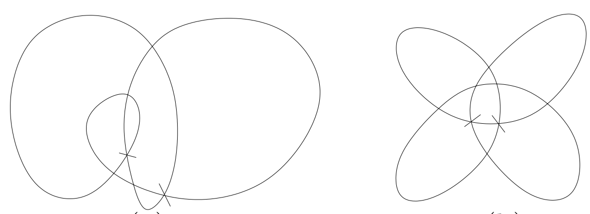


Figure 16: How to get rid of pseudo-eggs

Lemma 6.4 *Consider a curve $\gamma \in \mathcal{L}_I^{(1)} \setminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E}^+)$. For every direct loop (t_0, t_1) of γ there is a disjoint direct loop (t_2, t_3). For every reverse loop (t_0, t_1) of γ there is a disjoint (reverse or direct) loop (t_2, t_3). Furthermore, given two non-disjoint direct loops (t_0, t_1) and (t_2, t_3), γ admits a loop (t_4, t_5) which is disjoint from both (t_0, t_1) and (t_2, t_3).*

The second part does not always hold if (t_0, t_1) and (t_2, t_3) are reverse: see Figure 17, (a). In the last claim, the new loop in case (a) may be reverse: see Figure 17, (b).



(a) (b)

Figure 17: Non-disjoint loops (t_0, t_1) and (t_2, t_3).

Proof: Assume $\gamma \in \mathcal{L}_I^{(1)}$ so that every self-intersection is transversal. Let (t_0, t_1) be a loop. It $(t_1, t_0 + 1)$ is a loop then γ is a star. Otherwise, take $t_* = t_1 + \epsilon$ ($\epsilon > 0$, ϵ small). As in Lemma 6.2, let t_3 be the smallest $t > t_*$ such that $\gamma|_{[t_*, t]}$ is not injective and let $t_2 \in [t_*, t_3)$ be such that $\gamma(t_2) = \gamma(t_3)$, so that (t_2, t_3) is a loop. Since $(t_1, t_0 + 1)$ is not a loop, $t_3 < t_0 + 1$ and (t_2, t_3) is disjoint from (t_0, t_1), as required.

Let (t_0, t_1), (t_2, t_3) be two non-disjoint loops. The case $t_0 = t_2$, $t_1 = t_3$ was discussed in the previous paragraph; $t_2 = t_1$, $t_3 = t_0 + 1$ implies that γ is a star. We may therefore assume $t_0 < t_2 < t_1 < t_3 < t_0 + 1$, as in Figure 18.

Notice that $(\mathfrak{F}_\gamma(t_0))^{-1} \mathfrak{F}_\gamma(t_3) \in \mathcal{J}_2$. If the restriction $\gamma|_{[t_3, t_0+1]}$ is convex then γ is a star (Figure 18 (b)). If the restriction is injective but not convex then γ is either an egg or a pseudo-egg (Figure 18 (c)). Finally, if the above restriction is not injective then a new loop disjoint from (t_0, t_1) and (t_2, t_3) exists (Figure 18 (d)). ■

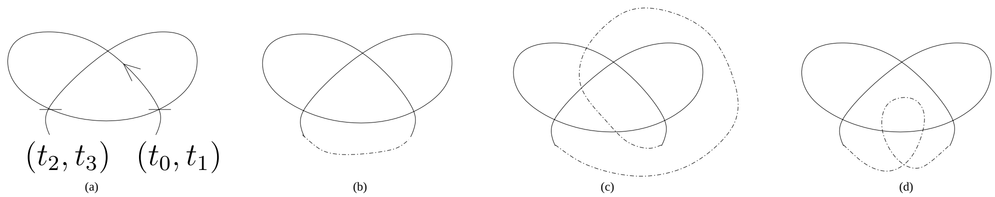


Figure 18: Two non-disjoint loops.

7 Proof of the main results

Lemma 7.1 *The connected components of $\mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$ are simply connected.*

Proof: Let $f : \mathbb{S}^1 \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$ be a continuous map. We prove that f is homotopic to a point. By Lemma 6.1 and transversality we may assume $f : \mathbb{S}^1 \to \mathcal{L}_I^{(2)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E})$ and that $f(p) \in \mathcal{L}_I^{(1)}$ except for finitely many points $p_1, \ldots, p_M \in \mathbb{S}^1$. In case $M \leq 2$ add a few points to the list to guarantee $M \geq 3$. Reparametrize so that $p_k = \frac{k}{M} \in \mathbb{S}^1$. By Lemma 6.2, for every $p \in \mathbb{S}^1$ the curve $f(p)$ has a direct loop. Let $V_k = (\frac{k-1}{M}, \frac{k+1}{M})$ and use a direct loop of $f(\frac{k}{M})$ to define $t_{0,k}, t_{1,k} : V_k \to \mathbb{S}^1$. This is an open cover, but the loops are probably not disjoint.

Let $U_k = (\frac{3k-1}{3M}, \frac{3k+1}{3M})$. Let $p_\star = \frac{2k+1}{2M}$, $\gamma_\star = f(p_\star) \in \mathcal{L}_I^{(1)}$, $t_0 = t_{0,k}(p_\star)$, $t_1 = t_{1,k}(p_\star)$, $t_2 = t_{0,k+1}(p_\star)$ and $t_3 = t_{1,k+1}(p_\star)$. If the loops (t_0, t_1) and (t_2, t_3) of $\gamma_\star$ are equal or disjoint, set $U_{k+1/2} = (\frac{k}{M}, \frac{k+1}{M})$ and $t_{*,k+1/2}(p) = t_{*,k}(p)$. If these two loops are not disjoint, use Lemma 6.4 to obtain two loops (t_4, t_5) and (t_6, t_7) of $\gamma_\star$ such that:

1. (t_4, t_5) is disjoint from (t_0, t_1);
2. (t_6, t_7) is disjoint from (t_2, t_3);
3. (t_4, t_5) and (t_6, t_7) are either equal or disjoint.

Set $U_{k+1/3} = (\frac{k}{M}, \frac{3k+2}{3M})$ and $U_{k+2/3} = (\frac{3k+1}{3M}, \frac{k+1}{M})$; use (t_4, t_5) and (t_6, t_7) to define $(t_{0,k+1/3}, t_{1,k+1/3})$ and $(t_{0,k+2/3}, t_{1,k+2/3})$, respectively. Thus, f admits a cover by disjoint loops. From Lemma 4.5, f is homotopic to $\nu_2 * f$.

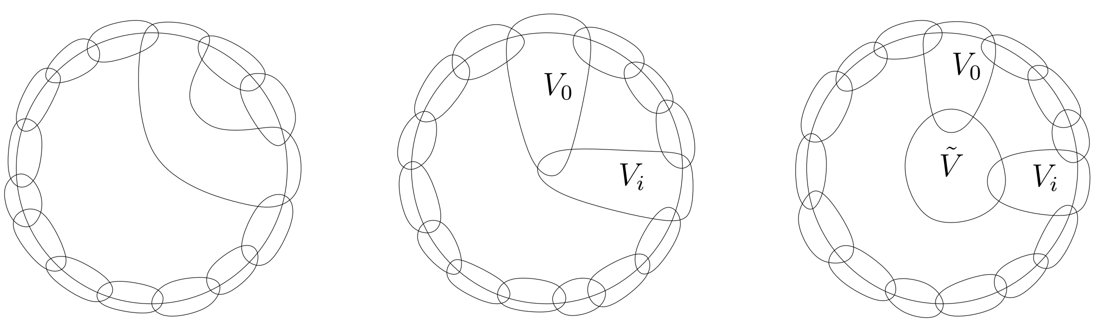
On the other hand, each connected component of $\mathcal{I}_I$ is simply connected and therefore f is homotopic to a point in $\mathcal{I}_I$. By Proposition 2.1 f is homotopic to a point in $\mathcal{L}_I$. The homotopy can be constructed as follows: consider an extension $\hat{f} : \mathbb{D}^2 \to \mathcal{I}_I$ of f; use $\Delta^\sharp$ to pass from f to $F_N \circ f$ and complete the homotopy with $F_N \circ \hat{f}$. For sufficiently large N, this homotopy remains in $\mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$. ■

Theorem 2 *Each connected component of $\mathcal{L}_I$ is simply connected.*

Proof: It is well known that $\mathcal{L}_{-1,c}$ is contractible. Since $\mathcal{T}_0 \subset \mathcal{L}_{+1}$, we just proved that in Lemma 7.1 that $\mathcal{L}_{-1,n}$ is also simply connected. Finally, we use Seifert-Van Kampen to compute $\pi_1(\mathcal{L}_{+1})$. Let $A = \mathcal{L}_{+1} \smallsetminus \mathcal{T}_0$ and fatten $\mathcal{T}_0$ a bit to obtain an open set B: since $\mathcal{T}_0$ is contractible and $\mathcal{T}_1 = \partial\mathcal{T}_0$ is a connected submanifold of codimension 1, B is simply connected and $A \cap B$ is connected. Thus, $\mathcal{L}_{+1} = A \cup B$ is also simply connected. ■

Lemma 7.2 *Let $\gamma \in \mathcal{L}_I^{(1)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E}^+)$. Let $f : \mathbb{D}^2 \to \mathcal{L}_I \smallsetminus \mathcal{T}_0$ be the function constant equal to γ. Consider a disjoint cover by loops $\mathcal{C}$ of $f|_{\mathbb{S}^1}$ without two consecutive reverse loops. Then there exists $\tilde{f} : \mathbb{D}^2 \to \mathcal{L}_I \smallsetminus \mathcal{T}_0$ homotopic to f with fixed boundary and an extension $\tilde{\mathcal{C}}$ of $\mathcal{C}$ to a disjoint cover by loops of $\tilde{f}$.*

Proof: Without loss of generality the open sets in the cover are intervals. If two neighboring intervals have identical loops, fuse them; may therefore assume without loss of generality that $\mathbb{S}^1$ is covered by a cycle of loops, where two adjacent loops are disjoint and no two adjacent loops are both reverse. More: if two non-adjacent intervals use the same loop, the corresponding open set can be enlarged to cross the disk (Figure 19,(a)). We may therefore assume all loops to be distinct.



(a) (b) (c)

Figure 19: Long cycles can be decomposed into short ones.

If the cycle 7 or more loops, it can be decomposed into cycles of size 6 or less. Indeed, consider a cycle of length $n > 6$; let ℓ_0 be a direct loop in the cycle and let $\ell_1, \ell_2, \ell_3, \ell_4$ be its neighbors (numbered clockwise). Let $i = 3$ or 4 such that ℓ_i is direct. If ℓ_0 and ℓ_i are disjoint, extend the open sets V_0 and V_i to intersect in the center, thus subdividing the original cycle into one of length $i+1$ and another of length $n-i+1$ (Figure 19,(b)). If ℓ_0 and ℓ_i are not disjoint, use Lemma 6.4 to obtain a loop $\tilde{\ell}$ disjoint from both: introduce an open set $\tilde{V}$ with associated loop $\tilde{\ell}$ in the middle of the disk, intersecting V_0 and V_i only among the originally defined loops. This subdivides the original cycle into two cycles of lengths $i+2$ and $n-i+2$ (Figure 19,(c)).

Cycles of 2 or 3 loops have all loops disjoint and therefore admit a disjoint cover as they are, no homotopy needed. We are left with considering cycles of length 4, 5 or 6.

From now on we argue by contradiction, searching for the shortest counterexample. If two non-adjacent loops are disjoint, not both reverse, construction (b) in Figure 19 can be applied, decomposing our cycle into two shorter ones. We may therefore assume that non-adjacent loops not both reverse are not disjoint.

We first consider the case $n = 6$. If two opposite loops ℓ_i and ℓ_{i+3} are both direct then the construction in Figure 19 (b) or (c) can be applied, decomposing our cycle. We may therefore assume that ℓ_i is direct for i even and reverse for i odd. Consider the intervals $I_i = [t_{0,i}, t_{1,i}] \subset \mathbb{S}^1$: I_0, I_2 and I_4 are pairwise neither disjoint not nested (one contained in the other). There are, up to reparametrization and permutation, only two possibilities: $I_0 = [0, 3/6]$, $I_2 = [1/6, 4/6]$, $I_4 = [-1/6, 2/6]$ or $I_0 = [0, 3/6]$, $I_2 = [2/6, 5/6]$, $I_4 = [-2/6, 1/6]$. In either case, I_3 must be disjoint from both I_2 and I_4 and neither disjoint nor nested with I_0, a contradiction.

Consider now the case $n = 5$, with loops $\ell_0, \ldots, \ell_4$. We may assume without loss of generality that ℓ_0, ℓ_2 and ℓ_3 are direct loops, ℓ_2 and ℓ_3 disjoint but ℓ_0 not disjoint from either. We may again assume that $I_0 = [1/10, 4/10]$, $I_2 = [0, 2/10]$, $I_3 = [3/10, 5/10]$ and that the image of $[0, 5/10]$ under γ is as in Figure 20. Pull the loop ℓ_0 (or, in other words, apply $\Delta^\sharp(s, \gamma, 1/10 - \epsilon, 4/10 + \epsilon)$) in the center of the disk $\mathbb{D}^2$ to define $\tilde{f}$. The loops ℓ_0, ℓ_2, ℓ_3 survive in $\mathbb{D}^2$ and near the center of $\mathbb{D}^2$, ℓ_0 becomes disjoint from ℓ_2, ℓ_3. The loops ℓ_1 and ℓ_4 were not affected and remain disjoint from ℓ_0. We therefore have the disjoint cover in Figure 20.

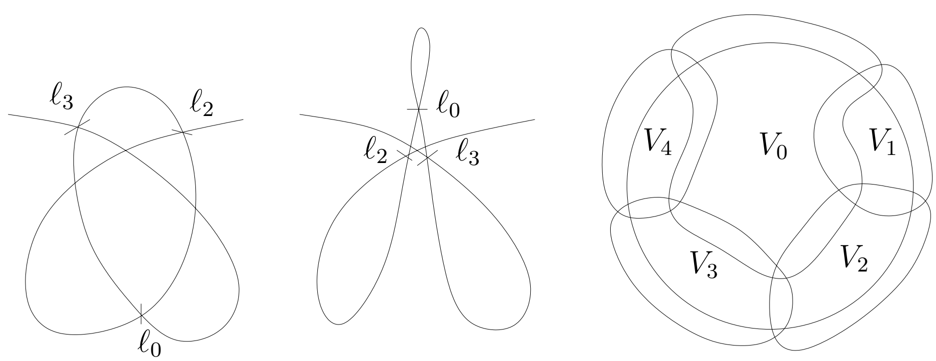


Figure 20: The case $n = 5$

Finally, for $n = 4$, we may assume that ℓ_0 and ℓ_2 are direct. As in Figure 21, pulling either ℓ_0 or ℓ_2 (or both) makes them disjoint. The disjoint cover in Figure 21 completes the proof. ■

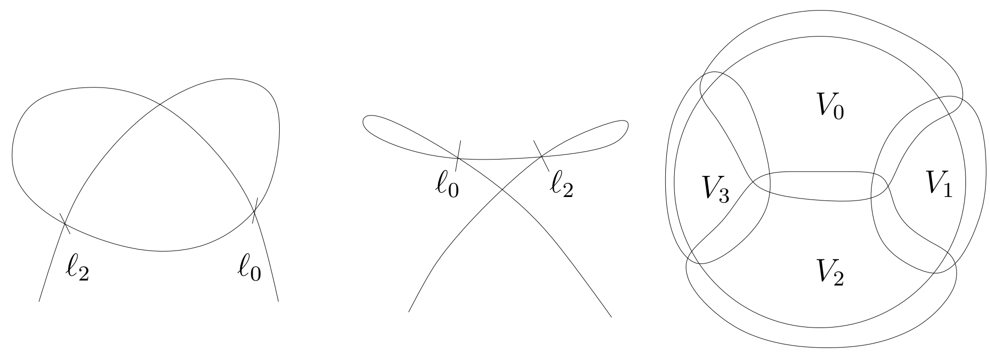


Figure 21: The case $n = 4$

Lemma 7.3 *Every continuous map $f : \mathbb{S}^2 \to \mathcal{L}_I \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0)$ is homotopic to $\nu_2 * f$.*

Proof: It is enough to prove that f is homotopic to some $\tilde{f}$ which admits a disjoint cover. By Lemmas 6.1, 6.3 and transversality we may assume $f : \mathbb{S}^1 \to \mathcal{L}_I^{(3)} \smallsetminus (\mathcal{L}_{-1,c} \cup \mathcal{T}_0 \cup \mathcal{E}^+)$. Again by transversality, we may assume that there exists a finite triangulation of $\mathbb{S}^2$ such that if $f(p) \notin \mathcal{L}_I^{(2)}$ then p is a vertex of the triangulation and if $f(p) \notin \mathcal{L}_I^{(1)}$ then p belongs to an edge. By Lemma 6.2, there exist open neighborhoods V_i of the vertices and direct loops $t_{0,i}, t_{1,i} : V_i \to \mathbb{S}^1$. As in the proof of Lemma 7.1, there exists a disjoint cover of each edge by loops without two consecutive reverse loops. It remain to fill in the faces: this is precisely what Lemma 7.2 does. ■

Theorem 3 *We have $\pi_2(\mathcal{L}_{+1}) = \mathbb{Z}^2$, $\pi_2(\mathcal{L}_{-1,n}) = \mathbb{Z}$, $H^2(\mathcal{L}_{+1};\mathbb{Z}) = \mathbb{Z}^2$ and $H^2(\mathcal{L}_{-1,n};\mathbb{Z}) = \mathbb{Z}$.*

Proof: By the previous lemma, $f : \mathbb{S}^2 \to \mathcal{L}_{-1,n}$ is homotopic to a point in $\mathcal{L}_{-1,n}$ if and only if it is homotopic to a point in $\mathcal{I}_{-1,n}$. In other words, the inclusion $\mathcal{L}_{-1,n} \subset \mathcal{I}_{-1}$ induces an isomorphism between $\pi_2(\mathcal{L}_{-1,n})$ and $\pi_2(\mathcal{I}_{-1}) = \mathbb{Z}$. By Hurewicz theorem, inclusion also yields an isomorphism between $H^2(\mathcal{L}_{-1,n};\mathbb{Z})$ and $H^2(\mathcal{I}_{-1};\mathbb{Z}) = \mathbb{Z}$. In other words, $H^2(\mathcal{L}_{-1,n};\mathbb{Z})$ is generated by $\mathbf{x}$.

Similarly, $H^2(\mathcal{L}_{+1} \smallsetminus \mathcal{T}_0;\mathbb{Z})$ is generated by $\mathbf{x}$. Use the normal bundle to $\mathcal{T}_1$ to define open sets A and B, $\mathcal{T}_0 \subset A$, $\mathcal{L}_{+1} \smallsetminus \mathcal{T}_0 \subset B$ such that the above inclusions and $\mathcal{T}_1 \subset A \cap B$ are homotopy equivalences. Write the Mayer-Vietoris sequence (coefficients in $\mathbb{Z}$):

$$H^1(A) \oplus H^1(B) \to H^1(A \cap B) \to H^2(A \cup B) \to H^2(A) \oplus H^2(B) \to H^2(A \cap B)$$

We know that $H^1(A) = H^1(B) = 0$, $H^1(A \cap B) = \mathbb{Z}$, $H^2(A) = 0$, $H^2(B) = \mathbb{Z}$, $H^2(A \cap B) = 0$. Thus $H^2(A \cup B) = \mathbb{Z}^2$. In the proof of Lemma 5.1 we saw a geometric description of the universal cover and therefore of the generator of $H^1(A \cap B)$: appying the map $H^1(A \cap B) \to H^2(A \cup B)$ to this generator obtains $\mathbf{f}_2 \in H^2(\mathcal{L}_{+1})$, the intersection number with $\mathcal{F}_2 \subset \mathcal{T}_1$. Thus, $H^2(\mathcal{L}_{+1}$ is generated by $\mathbf{x}$ and $\mathbf{f}_2$. Again by Hurewicz theorem, $\pi_2(\mathcal{L}_{+1}) = \mathbb{Z}^2$ is generated by $\mathbf{g}_2$ and $\nu_2 * \mathbf{g}_2$. ■

8 Final remarks

It may be possible to carry further methods used in this paper to compute $H^*(\mathcal{L}_I)$, but the kind of case-by-case analysis in Section 6 would have to be replaced by something less accidental. We hope to do this in [3] to prove that the classes $\mathbf{x}^n$ and $\mathbf{f}_{2n}$ are generators of $H^*(\mathcal{L}_{\pm 1})$ and that $\mathcal{L}_{+1}$ and $\mathcal{L}_{-1,n}$ have the homotopy type of $\Omega\mathbb{S}^3 \vee \mathbb{S}^2 \vee \mathbb{S}^6 \vee \mathbb{S}^{10} \vee \cdots$ and $\Omega\mathbb{S}^3 \vee \mathbb{S}^4 \vee \mathbb{S}^8 \vee \mathbb{S}^{12} \vee \cdots$, respectively.

References

[1] Burde, G. and Zieschang, H., *Knots*, Walter de Gruyter, Berlin, 1985.

[2] Little, J. A., *Nondegenerate homotopies of curves on the unit 2-sphere*, J. Differential Geometry, 4, 339-348, 1970.

[3] Saldanha, N., *The homotopy type of spaces of locally convex curves in the sphere*, in preparation.

[4] Saldanha, N., *The cohomology of spaces of locally convex curves in the sphere — I*, preprint, arXiv:0905.2111v1.

[5] Saldanha, N. and Shapiro, B., *Spaces of locally convex curves in S^n and combinatorics of the group B^+_{n+1}*, Journal of Singularities, 4, 1-22, 2012.

[6] Shapiro, B. and Khesin, B., *Homotopy classification of nondegenerate quasiperiodic curves on the 2-sphere*, Publ. Inst. Math. (Beograd) 66(80), 127-156, 1999.

[7] Shapiro, M., *Topology of the space of nondegenerate curves*, Math. USSR, 57, 106-126, 1993.

Nicolau C. Saldanha, PUC-Rio
saldanha@puc-rio.br; http://www.mat.puc-rio.br/∼nicolau/
Departamento de Matemática, PUC-Rio
R. Marquês de S. Vicente 225, Rio de Janeiro, RJ 22453-900, Brazil