Annals of Mathematics, **162** (2005), 1187–1241

Curve shortening and the topology of closed geodesics on surfaces

By Sigurd B. Angenent*

Abstract

We study "flat knot types" of geodesics on compact surfaces M^2. For every flat knot type and any Riemannian metric g we introduce a Conley index associated with the curve shortening flow on the space of immersed curves on M^2. We conclude existence of closed geodesics with prescribed flat knot types, provided the associated Conley index is nontrivial.

1. Introduction

If M is a surface with a Riemannian metric g then closed geodesics on (M, g) are critical points of the length functional $L(\gamma) = \int |\gamma'(x)| dx$ defined on the space of unparametrized C^2 immersed curves with orientation, i.e. we consider closed geodesics to be elements of the space

$$\Omega = \mathrm{Imm}(S^1, M)/\mathrm{Diff}_+(S^1).$$

Here $\mathrm{Imm}(S^1, M) = \{\gamma \in C^2(S^1, M) \mid \gamma'(\xi) \neq 0 \text{ for all } \xi \in S^1\}$ and $\mathrm{Diff}_+(S^1)$ is the group of C^2 orientation preserving diffeomorphisms of $S^1 = \mathbb{R}/\mathbb{Z}$. (We will abuse notation freely, and use the same symbol γ to denote both a convenient parametrization in $C^2(S^1; M)$, and its corresponding equivalence class in Ω.)

The natural gradient flow of the length functional is given by curve shortening, i.e. by the evolution equation

$$\frac{\partial \gamma}{\partial t} = \frac{\partial^2 \gamma}{\partial s^2} = \nabla_T(T), \qquad T \stackrel{\text{def}}{=} \frac{\partial \gamma}{\partial s}. \tag{1}$$

In 1905 Poincaré [33] pointed out that geodesics on surfaces are immersed curves *without self-tangencies.* Similarly, different geodesics cannot be tangent – all their intersections must be transverse. This allows one to classify closed geodesics by their number of self-intersections, or their "flat knot type,"

*Supported by NSF through a grant from DMS, and by the NWO through grant NWO-600-61-410.

and to ask how many closed geodesics of a given “type” exist on a given surface (M, g). Our main observation here is that the curve shortening flow (1) is the right tool to deal with this question.

We formalize these notions in the following definitions (which are a special case of the theory described by Arnol’d in [13].)

Flat knots. A curve $\gamma \in \Omega$ is a flat knot if it has no self-tangencies. Two flat knots α and β are equivalent if there is a continuous family of flat knots $\{\gamma_\theta \mid 0 \leq \theta \leq 1\}$ with $\gamma_0 = \alpha$ and $\gamma_1 = \beta$.

Relative flat knots. For a given finite collection of immersed curves,

$$\Gamma = \{\gamma_1, \ldots, \gamma_N\} \subset \Omega,$$

we define a *flat knot relative to* Γ to be any $\gamma \in \Omega$ which has no self-tangencies, and which is transverse to all $\gamma_j \in \Gamma$. Two flat knots relative to Γ are equivalent if one can be deformed into the other through a family of flat knots relative to Γ.

Clearly equivalent flat knots have the same number of self-intersections since this number cannot change during a deformation through flat knots. The converse is not true: Flat knots with the same number of self-intersections need not be equivalent. See Figure 1. Similarly, two equivalent flat knots relative to $\Gamma = \{\gamma_1, \ldots, \gamma_N\}$ have the same number of self-intersections, and the same number of intersections with each γ_j.

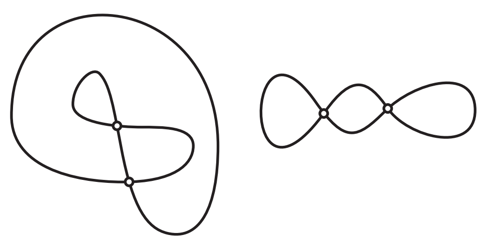

Figure 1: Two flat knots in $\mathbb{R}^2$ with two self-intersections

In this terminology any closed geodesic on a surface is a flat knot, and for given closed geodesics $\{\gamma_1, \ldots, \gamma_N\}$ any other closed geodesic is a flat knot relative to $\{\gamma_1, \ldots, \gamma_N\}$.

One can now ask the following question: Given a Riemannian metric g on a surface M, closed geodesics $\gamma_1, \ldots, \gamma_N$ for this metric, and a flat knot α relative to $\Gamma = \{\gamma_1, \ldots, \gamma_N\}$, *how many closed geodesics on* (M, g) *define flat knots relative to* Γ *which are equivalent to* α? In this paper we will use curve shortening to obtain a lower bound for the number of such closed geodesics which only depends on the relative flat knot α, and the linearization of the geodesic flow on (TM, g) along the given closed geodesics γ_j.

Our strategy for estimating the number of closed geodesics equivalent to a given relative flat knot α is to consider the set $\mathcal{B}_\alpha \subset \Omega$ of all flat knots relative to Γ which are equivalent to α. This set turns out to be *almost* an isolating block in the sense of Conley [17] for the curve shortening flow. We then define a Conley index $h(\mathcal{B}_\alpha)$ of $\mathcal{B}_\alpha$ and use standard variational arguments to conclude that nontriviality of the Conley index of a relative flat knot implies existence of a critical point for curve shortening in $\mathcal{B}_\alpha$.

To do all this we have to overcome a few obstacles.

First, the curve shortening flow is not a globally defined flow or even semiflow. Given any initial curve $\gamma(0) \in \Omega$ a solution $\gamma : [0, T) \to \Omega$ to curve shortening exists for a short time $T = T(\gamma_0) > 0$, but this solution often becomes singular in finite time. What helps us overcome this problem is that the set of initial curves $\gamma(0) \in \mathcal{B}_\alpha$ which are close to forming a singularity is attracting. Indeed, the existing analysis of the singularities of curve shortening in [24], [7], [25], [26], [32] shows that such singularities essentially only form when "a small loop in the curve $\gamma(t)$ contracts as $t \nearrow T(\gamma(0))$." A calculation involving the Gauss-Bonnet theorem shows that once a curve has a sufficiently small loop the area enclosed by this loop must decrease under curve shortening. This observation allows us to include the set of curves $\gamma \in \mathcal{B}_\alpha$ with a small loop in the exit set of the curve shortening flow. With this modification we can proceed as if the curve shortening flow were defined globally.

Second, $\mathcal{B}_\alpha$ is not a closed subset of Ω and its boundary may contain closed geodesics, i.e. critical points of curve shortening: such critical points are always multiple covers of shorter geodesics. To deal with this, one must analyze the curve shortening flow near multiple covers of closed geodesics. It turns out that all relevant information to our problem is contained in Poincaré's rotation number of a closed geodesic. In the end our Conley index $h(\mathcal{B}_\alpha)$ depends not only on the relative flat knot class $\mathcal{B}_\alpha$, but also on the rotation numbers of the given closed geodesics $\{\gamma_1, \ldots, \gamma_N\}$.

Finally, the space $\mathcal{B}_\alpha$ on which curve shortening is defined is not locally compact so that Conley's theory does not apply without modification. It turns out that the regularizing effect of curve shortening provides an adequate substitute for the absence of local compactness of $\mathcal{B}_\alpha$.

After resolving these issues one merely has to compute the Conley index of any relative flat knot type to estimate the number of closed geodesics of that type. To describe the results we need to discuss satellites and Poincaré's rotation number.

1.1. *Satellites.* Let $\alpha \in \Omega$ be given, and let $\alpha : \mathbb{R}/\mathbb{Z} \to M$ also denote a constant speed parametrization of α. Choose a unit normal $\mathbf{N}$ along α, and consider the curve $\alpha_\epsilon : \mathbb{R}/\mathbb{Z} \to M$ given by

$$\alpha_\epsilon(t) = \exp_{\alpha(qt)}\bigl(\epsilon \sin(2\pi pt)\mathbf{N}(qt)\bigr)$$

where $\frac{p}{q}$ is a fraction in lowest terms. When $\epsilon = 0$, α_ϵ is a q-fold cover of α. For sufficiently small $\epsilon \neq 0$ the α_ϵ are flat knots relative to α. Any flat knot relative to α equivalent to α_ϵ is by definition a (p, q)-satellite of α.

Poincaré [33] observed that a (p, q)-satellite of a simple closed curve α has $2p$ intersections with α and $p(q-1)$ self-intersections. See also Lemma 2.1.

1.2. *Poincaré's rotation number.* Let $\gamma(s)$ be an arc-length parametrization of a closed geodesic of length $L > 0$ on (M, g). Thus $\gamma(s+L) \equiv \gamma(s)$, and $T = \gamma'(s)$ satisfies $\nabla_T T = 0$. Jacobi fields are solutions of the second order ODE

$$\frac{d^2y}{ds^2} + K(\gamma(s))y(s) = 0, \tag{2}$$

where $K : M \to \mathbb{R}$ is the Gaussian curvature of (M, g).

Let $y : \mathbb{R} \to \mathbb{R}$ be any Jacobi field, and label the zeroes of y in increasing order

$$\ldots < s_{-2} < s_{-1} < s_0 < s_1 < s_2 < \ldots$$

with $(-1)^n y'(s_n) > 0$. Using the Sturm oscillation theorems one can then show that the limit

$$\omega(\gamma) = \lim_{n\to\infty} \frac{s_{2n}}{nL}$$

exists and is independent of the chosen Jacobi field y. We call this number the Poincaré rotation number of the geodesic γ. If there is a Jacobi field with only finitely many zeroes then the oscillation theorems again imply that $y(s)$ has either one or no zeroes $s \in \mathbb{R}$. In this case we say the rotation number is infinite.

For an alternative definition we observe that if $y(s)$ is a Jacobi field then $y(s)$ and $y'(s)$ cannot vanish simultaneously. Thus one can consider

$$\rho(\gamma) = \lim_{s\to\infty} \frac{L}{2\pi s} \arg\{y(s) + iy'(s)\}.$$

Again it turns out that this limit exists and is independent of the particular choice of Jacobi field y. Moreover one has

$$\rho = \frac{1}{\omega}.$$

We call ρ the inverse rotation number of γ. See [27] where the much more complicated case of quasi-periodic potentials is treated. The inverse rotation number ρ is analogous to the "amount of rotation" of a periodic orbit of a twist map introduced by Mather in [30].

1.3. *Allowable metrics for a given relative flat knot and the nonresonance condition.* Let $\Gamma = \{\gamma_1, \ldots, \gamma_N\} \subset \Omega$ be a collection of curves with no mutual or self-tangencies, and denote by $\mathcal{M}_\Gamma$ the space of $C^{2,\mu}$ Riemannian metrics g on M for which the $\gamma_i \in \Gamma$ are geodesics (thus the metric has continuous derivatives of second order which are Hölder continuous of some exponent $\mu \in (0,1)$). When written out in coordinates one sees that this condition is quadratic in the components g_{ij} and $\partial_i g_{jk}$ of the metric and its derivatives. Thus $\mathcal{M}_\Gamma$ is a closed subspace of the space of $C^{2,\mu}$ metrics on $\mathcal{M}$.

If $\alpha \in \Omega$ is a flat knot rel Γ then it may happen that α is a (p_1, q_1) satellite of, say, γ_1. In this case the rotation number of γ_1 will affect the number of closed geodesics of flat knot type α rel Γ. To see this, consider a family of metrics $\{g_\lambda \mid \lambda \in \mathbb{R}\} \subset \mathcal{M}_\gamma$ for which the inverse rotation number $\rho(\gamma; g_\lambda)$ is less than p_1/q_1 for negative λ and more than p_1/q_1 for positive λ. Then, as λ increases from negative to positive, a bifurcation takes place in which generically two (p_1, q_1) satellites of γ_1 are created. These bifurcations appear as resonances in the Birkhoff normal form of the geodesic flow on the unit tangent bundle near the lift of γ. This is described by Poincaré in [33, §6, p. 261]. See also [14, Appendix 7D,F].

In studying the closed geodesics of flat knot type α rel Γ we will therefore exclude those metrics for which a bifurcation can take place. To be precise, given α we order the γ_i so that α is a (p_i, q_i) satellite of γ_i, if $1 \le i \le m$, but not a satellite of γ_i for $m < i \le N$. We then impose the nonresonance condition

$$\rho(\gamma_i) \neq \frac{p_i}{q_i} \text{ for } i \in \{1, \dots, m\}. \tag{3}$$

The metrics $g \in \mathcal{M}_\Gamma$ which satisfy this condition can be separated into 2^m distinct classes. For any subset $I \subset \{1, \dots, m\}$ we define $\mathcal{M}_\Gamma(\alpha; I)$ to be the set of all metrics $g \in \mathcal{M}_\Gamma$ such that the inverse rotation numbers $\rho(\gamma_1)$, ... , $\rho(\gamma_m)$ satisfy

$$\rho(\gamma_i) < \frac{p_i}{q_i} \text{ if } i \in I \text{ and } \rho(\gamma_i) > \frac{p_i}{q_i} \text{ if } i \notin I. \tag{4}$$

For each $I \subset \{1, \dots, m\}$ we define in Section 6 a Conley index h^I. This is done by choosing a metric $g \in \mathcal{M}_\Gamma(\alpha; I)$, suitably modifying the set $\mathcal{B}_\alpha \subset \Omega$ and its exit set for the curve shortening flow, according to the choice of $I \subset \{1, \dots, m\}$ and then finally setting h^I equal to the homotopy type of the modified $\mathcal{B}_\alpha$ with its exit set collapsed to a point. Thus the index we define is the homotopy type of a topological space with a distinguished point. We show that the resulting index h^I does not depend on the choice of metric $g \in \mathcal{M}(\alpha; I)$, and also that the index h^I does not change if one replaces α by an equivalent flat knot rel Γ.

Using rather standard variational methods we then show in §7:

THEOREM 1.1. *If $g \in \mathcal{M}_\Gamma(\alpha; I)$ and if the index h^I is nontrivial, then the metric g has at least one closed geodesic of flat knot type α* rel Γ.

Using more standard variational arguments one could then improve on this and show that there are at least $n-1$ closed geodesics of type α rel Γ, where n is the Lyusternik-Schnirelman category of the pointed topological space h^I. We do not use this result here and omit the proof.

Computation of the index h^I for an arbitrary flat knot α rel Γ may be difficult. It is simplified somewhat by the independence of h^I from the metric $g \in \mathcal{M}_\Gamma(\alpha; I)$. In addition we have a long exact sequence which relates the homologies of the different indices one gets by varying I.

THEOREM 1.2. *Let $\varnothing \subset J \subset I \subset \{1, \ldots, m\}$ with $J \neq I$. Then there is a long exact sequence*

$$\ldots H_{l+1}(h^I) \xrightarrow{\partial_*} H_l(\mathcal{A}_J^I) \longrightarrow H_l(h^J) \longrightarrow H_l(h^I) \xrightarrow{\partial_*} H_{l-1}(\mathcal{A}_J^I) \ldots \tag{5}$$

where

$$\mathcal{A}_J^I = \bigvee_{k \in I \setminus J} \left\{ \frac{S^1 \times S^{2p_k - 1}}{S^1 \times \{pt\}} \right\}.$$

This immediately implies

THEOREM 1.3. *If $J \subset I$ with $J \neq I$ then h^I and h^J cannot both be trivial.*

One may regard this as a global bifurcation theorem. If for some choice of rotation numbers I and some choice of metric $g \in \mathcal{M}_\Gamma(\alpha; I)$ there are no closed geodesics of type α rel Γ, then the index h^I is trivial. By increasing one or more of the rotation numbers (i.e. increasing I to J), or by decreasing some of the rotation numbers (i.e. decreasing I to J) the index h^I becomes nontrivial, and a closed geodesic of type α rel Γ must exist *for any metric $g \in \mathcal{M}_\Gamma(\alpha; J)$*.

When applied to the case where $M = S^2$ and Γ consists of one simple closed curve γ this gives us the following result.

THEOREM 1.4. *Let g be a $C^{2,\mu}$ metric on M with a simple closed geodesic $\gamma \in \Omega$. Let $\rho = \rho(\gamma, g)$ be the inverse rotation number of γ.*

If $\rho > 1$ then for each $\frac{p}{q} \in (1, \rho)$ there is a closed geodesic $\gamma_{p/q}$ on (M, g) which is a (p, q) satellite of γ.

Similarly, if $\rho < 1$ then for each $\frac{p}{q} \in (\rho, 1)$ there is a closed geodesic $\gamma_{p/q}$ on (M, g) which is a (p, q) satellite of γ.

In both cases the geodesic $\gamma_{p/q}$ intersects the given simple closed geodesic γ exactly $2p$ times, and $\gamma_{p/q}$ intersects itself exactly $p(q-1)$ times.

Acknowledgements. The work in this paper was inspired by a question of Hofer (Oberwollfach, 1993) who asked me if one could apply the Floer homology construction to curve shortening, and which results could be obtained in this way. This turned out to be a very fruitful question, even though in the

end curve shortening appears to be sufficiently well behaved to use the Conley index instead of Floer's approach.

The paper was finished during my sabattical at the University of Leiden. It is a pleasure to thank Rob van der Vorst, Bert Peletier and Sjoerd Verduyn Lunel for their hospitality.

Contents

1. Introduction
2. Flat knots
3. Curve shortening
4. Curve shortening near a closed geodesic
5. Loops
6. Definition of the Conley index of a flat knot
7. Existence theorems for closed geodesics
8. Appendices
References

2. Flat knots

2.1. *The space of immersed curves.* The space of immersed curves $\Omega = \mathrm{Imm}(S^1,\ M)/\mathrm{Diff}_+\left(S^1\right)$ is locally homeomorphic to $C^2(\mathbb{R}/\mathbb{Z})$. The homeomorphisms are given by the following charts. Let $\gamma \in \Omega$ be a given immersed curve. Choose a C^2 parametrization $\gamma : \mathbb{R}/\mathbb{Z} \to M$ of this curve and extend it to a C^2 local diffeomorphism $\sigma : (\mathbb{R}/\mathbb{Z}) \times (-r, r) \to M$ for some $r > 0$. Then for any C^1 small function $u \in C^2(\mathbb{R}/\mathbb{Z})$ the curve

$$\gamma_u(x) = \sigma(x, u(x)) \tag{6}$$

is an immersed C^2 curve. Let $\mathcal{U}_r = \{u \in C^2(\mathbb{R}/\mathbb{Z}) : |u(x)| < r\}$. For sufficiently small $r > 0$ the map $\Phi : u \in \mathcal{U}_r \mapsto \gamma_u \in \Omega$ is a homeomorphism of $\mathcal{U}_r$ onto a small neighborhood $\Phi(\mathcal{U}_r)$ of γ. The open sets $\Phi(\mathcal{U}_r)$ which one gets by varying the curve γ cover Ω, and hence Ω is a topological Banach manifold with model $C^2(\mathbb{R}/\mathbb{Z})$.

A natural choice for the local diffeomorphism σ would be

$$\sigma(x, u) = \exp_{\gamma(x)}(u\mathbf{N}(x))$$

where $\mathbf{N}$ is a unit normal vector field for the curve γ. We avoid this choice of σ since it uses too many derivatives. For σ to be C^2 one would want the normal to be C^2, so the curve would have to be C^3; one would also want the exponential map to be C^2, which requires the Christoffel symbols to have two derivatives, and so the metric g would have to be C^3.

For future reference we observe that if the curve γ is $C^{2,\mu}$ then one can also choose the diffeomorphism σ to be $C^{2,\mu}$.

2.2. *Covers.* For any $\gamma \in \Omega$ and any nonzero integer q we define $q \cdot \gamma$ to be the q-fold cover of γ, i.e. the curve with parametrization

$$(q \cdot \gamma)(t) = \gamma(qt), \quad t \in \mathbb{R}/\mathbb{Z},$$

where $\gamma : \mathbb{R}/\mathbb{Z} \to M$ is a parametrization of γ. Thus $(-1) \cdot \gamma$ is the curve γ with its orientation reversed.

A curve $\gamma \in \Omega$ will be called *primitive* if it is not a multiple cover of some other curve, i.e. if there are no $q \geq 2$ and $\gamma_0 \in \Omega$ with $\gamma = q \cdot \gamma_0$.

2.3. *Flat knots.* Let $\gamma_1, \ldots, \gamma_N$ be a collection of primitive immersed curves in M. Define

$$\Delta(\gamma_1, \ldots, \gamma_N) = \left\{ \gamma \in \Omega \;\middle|\; \begin{array}{l} \gamma \text{ has a self-tangency or a} \\ \text{tangency with one of the} \\ \gamma_i \end{array} \right\} \tag{7}$$

and

$$\Delta = \{\gamma \in \Omega \mid \gamma \text{ has a self-tangency}\}. \tag{8}$$

Then Δ and $\Delta(\gamma_1, \ldots, \gamma_N)$ are closed subsets of Ω, and their complements $\Omega \setminus \Delta$ and $\Omega \setminus \Delta(\gamma_1, \ldots, \gamma_N)$ consist of flat knots, and flat knots relative to $(\gamma_1, \ldots, \gamma_N)$, respectively. Two such flat knots are equivalent if and only if they lie in the same component of $\Omega \setminus \Delta$ or $\Omega \setminus \Delta(\gamma_1, \ldots, \gamma_N)$.

2.4. *Flat knots as knots in the projective tangent bundle.* Let $\mathbb{P}TM$ be the *projective tangent bundle* of M, i.e. $\mathbb{P}TM$ is the bundle obtained from the unit tangent bundle

$$T^1(M) = \{(p, v) \in T(M) \mid g(v, v) = 1\}$$

by identification of all antipodal vectors (x, v) and $(x, -v)$. The projective tangent bundle is a contact manifold. If we denote the bundle projection by $\pi : \mathbb{P}TM \to M$, then the contact plane $L_{(x,\pm v)} \subset T(\mathbb{P}TM)$ at a point $(x, \pm v) \in \mathbb{P}TM$ consists of those vectors $\xi \in T(\mathbb{P}TM)$ for which $d\pi(\xi)$ is a multiple of v. Each contact plane $L_{(x,\pm v)}$ contains a nonzero vector ϑ with $d\pi(\vartheta) = 0$ (ϑ corresponds to infinitesimal rotation of the unit vector $\pm v$ in the tangent space T_xM, while the base point x remains fixed).

Any $\gamma \in \Omega$ defines a C^1 immersed curve $\hat{\gamma}$ in the projective tangent bundle $\mathbb{P}TM$ with parametrization $\hat{\gamma}(s) = (\gamma(s), \pm\gamma'(s))$, where $\gamma(s)$ is an arc length parametrization of γ. We call $\hat{\gamma}$ the lift of γ.

An immersed curve $\tilde{\gamma}$ in $\mathbb{P}TM$ is the lift of some $\gamma \in \Omega$ if and only if $\tilde{\gamma}$ is everywhere tangent to the contact planes, and nowhere tangent to the special direction ϑ in the contact planes.

Self-tangencies of $\gamma \in \Omega$ correspond to self-intersections of its lift $\hat{\gamma} \subset \mathbb{P}TM$. Thus an immersed curve $\gamma \in \Omega$ is a flat knot exactly when its lift $\hat{\gamma}$ is a knot in the three manifold $\mathbb{P}TM$. If two curves $\gamma_1, \gamma_2 \in \Omega$ define equivalent flat knots then one can be deformed into the other through flat knots. By lifting the deformation we see that $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are equivalent knots in $\mathbb{P}TM$.

2.5. *Intersections.* If $\alpha \in \Omega \setminus \Delta(\gamma_1, \ldots, \gamma_n)$ then α is transverse to each of the γ_i. Hence the number of intersections in $\alpha \cap \gamma_i$ is well defined. This only depends on the flat knot type of α relative to $\gamma_1, \ldots, \gamma_n$.

If $\alpha \in \Omega \setminus \Delta$ then α only has transverse self-intersections, so their number is well defined by $\#\alpha \cap \alpha = \#\{0 \leq x < x' < 1 \mid \alpha(x) = \alpha(x')\}$. From a drawing of α they are easily counted. An $\alpha \in \Omega \setminus \Delta$ can only have double points, triple points, etc. (see Figure 2). If α only has double points (a generic property) then their number is the number of self-intersections. Otherwise one must count the number of geometric self-intersections where a k-tuple point counts for $\binom{k}{2}$ self-intersections. Again this number only depends on the flat knot type of $\alpha \in \Omega \setminus \Delta$.

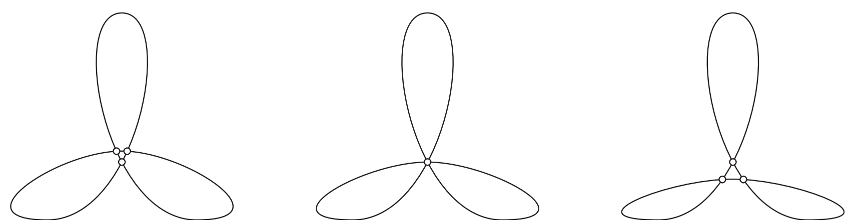

Figure 2: Equivalent flat knots with 3 self-intersections.

2.6. *Nontransverse crossings of curves.* If $\gamma_1, \gamma_2 \in \Omega$ are not necessarily transverse then we define the number of crossings of γ_1 and γ_2 to be

$$\text{Cross}(\gamma_1, \gamma_2) = \sup_{\gamma_i \in \mathcal{U}_i} \inf \left\{ \#(\gamma_1' \cap \gamma_2') \;\middle|\; \begin{array}{c} \gamma_1' \in \mathcal{U}_1,\ \gamma_2' \in \mathcal{U}_2 \\ \gamma_1' \pitchfork \gamma_2' \end{array} \right\} \tag{9}$$

where the supremum is taken over all pairs of open neighborhoods $\mathcal{U}_i \subset \Omega$ of γ_i. Thus $\text{Cross}(\gamma_1, \gamma_2)$ is the smallest number of intersections γ_1 and γ_2 can have if one perturbs them slightly to become transverse.

The number of self-crossings $\text{Cross}(\gamma, \gamma)$ is defined in a similar way.

Clearly $\text{Cross}(\gamma_1, \gamma_2)$ is a lower semicontinuous function on $\Omega \times \Omega$.

2.7. *Satellites.* We first describe the local model of a satellite of a primitive flat knot $\gamma \in \Omega \setminus \Delta$ and then transplant the local model to primitive flat knots on any surface.

Let $q \geq 1$ be an integer, and let $u \in C^2(\mathbb{R}/q\mathbb{Z})$ be a function for which

$$\text{all zeroes of } u \text{ are simple} \tag{10}$$

and

$$\text{all zeroes of } v_k(x) \stackrel{\text{def}}{=} u(x) - u(x-k) \text{ are simple for } k = 1, 2, \cdots, q-1. \tag{11}$$

Consider the curve α_u in the cylinder $\Gamma = (\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$, parametrized by

$$\alpha_u : \mathbb{R}/q\mathbb{Z} \to \Gamma, \quad \alpha_u(x) = (x, u(x)). \tag{12}$$

The conditions (10) and (11) imply that α_u is a flat knot relative to α_0, where $\alpha_0 = (\mathbb{R}/\mathbb{Z}) \times \{0\}$ is the zero section (i.e., the curve corresponding to $u(x) \equiv 0$).

Now consider a primitive flat knot $\gamma \in \Omega \backslash \Delta$. Denote by $\gamma : \mathbb{R}/\mathbb{Z} \to M$ any parametrization, and choose a local diffeomorphism $\sigma : \mathbb{R}/\mathbb{Z} \times (-r, r) \to M$ with $\gamma(x) = \sigma(x, 0)$. As in §2.1 we then identify any curve γ_u which is C^1 close to γ with a function $u \in C^2(\mathbb{R}/\mathbb{Z})$ via (6).

If $u \in C^2(\mathbb{R}/q\mathbb{Z})$ then the curve defined by

$$\alpha_{\varepsilon,u}(x) = \sigma(x, \varepsilon u(x)) \tag{13}$$

is a flat knot relative to γ. For given $u \in C^2(\mathbb{R}/q\mathbb{Z})$ and small enough $\varepsilon > 0$ the $\alpha_{\varepsilon,u}$ all define the same relative flat knot.

By definition, a curve $\alpha \in \Omega \setminus \Delta(\gamma)$ is a satellite of $\gamma \in \Omega \setminus \Delta$ if for some $u \in C^2(\mathbb{R}/q\mathbb{Z})$ it is isotopic relative to γ to all $\alpha_{\varepsilon,u}$ with $\varepsilon > 0$ sufficiently small.

To complete this definition we should specify the orientation of the satellite $\alpha_{\varepsilon,u}$. One can give $\alpha_{\varepsilon,u}$ as defined in (13) the same orientation as its base curve γ, or the opposite orientation. We will call both curves satellites of γ. In general the satellites $\alpha_{\varepsilon,u}$ and $-\alpha_{\varepsilon,u}$ can define different flat knots relative to γ or they can belong to the same relative flat knot class.

Example. Let γ be the equator on the standard two sphere $M = S^2$. Then any other great circle is a satellite of γ. Moreover, all these great circles with either orientation define the same flat knot relative to the equator. For example, if α is a great circle in a plane through the x-axis which makes an angle $\varphi \ll \pi/2$ with the xy-plane, then one can reverse its orientation by first rotating it through $\pi - 2\varphi$ around the x-axis, and then rotating it through π around the z-axis. Throughout this motion the curve remains transverse to the equator, so that α and $-\alpha$ indeed belong to the same component of $\Delta \setminus \Omega(\gamma)$. Below we will show that this example is exceptional.

As defined in the introduction, one obtains (p, q) satellites by setting

$$u(x) = \sin(2\pi \frac{p}{q} x). \tag{14}$$

Let $p \neq 0$, and let α be the (p, q) satellite of γ given by $u(x) = \epsilon \sin(2\pi \frac{p}{q} x)$. Then we can translate α along the base curve γ; i.e. we can consider the (p, q) satellites α_τ given by $u_\tau(x) = \epsilon \sin(2\pi \frac{p}{q}(x - \tau))$. By translating from $\tau = 0$ to $\tau = \frac{q}{2p}$ one finds an isotopy from α to the curve $\bar{\alpha}$ given by $\bar{u}(x) = \sin(2\pi \frac{-p}{q} x)$. Hence one can turn any (p, q) satellite into a $(-p, q)$ satellite, and we may therefore always assume that p is nonnegative.

We will denote the set of (p,q)-satellites of $\gamma \in \Omega$ by $\mathcal{B}_{p,q}(\gamma)$, always assuming that $p \geq 0$ and $q \geq 1$.

More precisely we will let $\mathcal{B}^+_{p,q}(\gamma)$ be the set of (p,q)-satellites of γ which have the same orientation as γ, and we let $\mathcal{B}^-_{p,q}(\gamma)$ be those (p,q) satellites with opposite orientation. With this notation we always have

$$\mathcal{B}_{p,q}(\zeta) = \mathcal{B}^+_{p,q}(\zeta) \cup \mathcal{B}^-_{p,q}(\zeta).$$

It is not *a priori* clear that all these classes are disjoint, but by counting the number of self-intersections of (p,q) satellites one can at least see that there are infinitely many disjoint $\mathcal{B}_{p,q}$'s.

LEMMA 2.1. *Let $\gamma \in \Omega \setminus \Delta$ be a flat knot with m self-intersections. Then any $\alpha \in \mathcal{B}_{p,q}(\gamma)$ has exactly $2p + 2mq$ intersections with ζ, and $p(q-1) + mq^2$ self-intersections.*

This was observed by Poincaré [33]. We include a proof for completeness' sake.

Proof. Intersections of α and γ are of two types. Each zero of $u(x)$ corresponds to an intersection of α and γ. At each self-intersection of γ the two intersecting strands of γ are accompanied by $2q$ strands of α which intersect γ in $2q$ points. Since $u(x)$ has $2p$ zeroes and γ has m self-intersections we get $2mq + 2p$ intersections of α and γ.

To count self-intersections one must count the intersections of the graph of $u(x) = \sin(2\pi\frac{p}{q}x)$ wrapped up on the cylinder $\Gamma = (\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$, i.e. the intersections of the graphs of $u_k(x) = u(x - 2k)$ $(k = 0, 1, \ldots, q-1)$ with $0 \leq x < 2\pi$. After some work one finds that these are arranged in $q - 1$ horizontal rows, each of which contains p intersections.

At each self-intersection of γ two strands of γ cross. If ε is small enough then $\alpha_{\varepsilon,u}$ is locally almost parallel to γ, so that any pair of crossing strands of γ is accompanied by a pair of q nearly parallel strands of α which cross each other. This way we get q^2 extra self-crossings of α and $2q$ extra crossings of γ with α per self-crossing of γ. □

LEMMA 2.2. *If $\mathcal{B}_{p,q}(\gamma) \cap \mathcal{B}_{r,s}(\gamma) \neq \varnothing$ then $p = r$ and $q = s$.*

Proof. If $\alpha \in \mathcal{B}_{p,q}$ has $2k$ intersections with γ and l self-intersections then

$$p(q-1) + mq^2 = l, \qquad p + mq = k.$$

Substitute $p = k - mq$ in the first equation to get

$$l = (k+m)q - k = mq + (q-1)k$$

from which one finds $q = \frac{k+l}{k+m}$. In particular, the numbers k, l and m determine p and q. □

The proof also shows that most satellites are not (p,q)-satellites for any (p,q). Indeed, given $\alpha \in \mathcal{B}_{p,q}(\gamma)$ one can modify it near one of its crossings with γ so as to increase the number k of intersections with γ arbitrarily without changing the number of self-intersections l, or m. Unless both $l=0$ and $m=0$, then for large enough k the fraction $\frac{k+l}{k+m}$ will not be an integer, so the modified curve can no longer be a (p,q) satellite. If both $l=m=0$ then both γ and its satellite α must be simple curves.

2.8. *(p,q) satellites along a simple closed curve on S^2.* In this section we consider the case in which $M=S^2$ and $\zeta \in \Omega$ is a simple closed curve. We will show that for all (p,q) except $p=q=1$ the classes $\mathcal{B}^{\pm}_{p,q}(\zeta)$ are different.

After applying a diffeomorphism we may assume that M is the unit sphere in $\mathbb{R}^3$ and that ζ is the equator, given by $z=0$.

To study curves in $\Omega \setminus \Delta(\zeta)$ it is useful to recall that one can identify the unit tangent bundle $T^1(S^2)$ of the 2-sphere with the group $\mathrm{SO}(3,\mathbb{R})$. Indeed, by definition,

$$T^1(S^2) = \{(\vec{x},\vec{\xi}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid |\vec{x}| = |\vec{\xi}| = 1, \vec{x} \perp \vec{\xi}\}$$

so that any unit tangent vector $(\vec{x},\vec{\xi}) \in T^1(S^2)$ determines the first two columns of an orthogonal matrix. The third column of this matrix is the cross product $\vec{x} \times \vec{\xi}$. The map

$$(\vec{x},\vec{\xi}) \in T^1(S^2) \mapsto (\vec{x},\vec{\xi},\vec{x}\times\vec{\xi}) \in \mathrm{SO}(3,\mathbb{R})$$

is a diffeomorphism, and from here on we will simply identify $T^1(S^2)$ and $\mathrm{SO}(3,\mathbb{R})$.

Let $\mathcal{U} \subset T^1(S^2)$ be the complement of the set of tangent vectors to ζ and $-\zeta$. One can describe $\mathcal{U}$ very conveniently using "Euler Angles". For the definition of these angles we refer to Figure 3. Any unit tangent vector $(\vec{x},\vec{\xi})$ defines an oriented great circle, parametrized by

$$X(t) = (\cos t)\vec{x} + (\sin t)\vec{\xi}.$$

Unless $(\vec{x},\vec{\xi})$ is a tangent vector of the equator $\pm\zeta$, the great circle through $(\vec{x},\vec{\xi})$ intersects the equator in two points. In one of these intersections the great circle crosses the equator from south to north. Let θ be the angle from the upward intersection to x, so that $X(-\theta)$ is the upward intersection point. We define ψ to be the angle between the plane through the great circle $\{X(t) \mid t \in \mathbb{R}\}$ and the xy-plane (so that $0 < \psi < \pi$). Finally we let ϕ be the angle along the equator ζ from the x-axis to the upward intersection point $X(-\theta)$.

If we denote the matrix corresponding to a rotation by an angle α around the x axis by $\mathrm{R}_x(\alpha)$, etc. then the relation between the Euler angles (θ,φ,ψ) and the unit tangent vector (x,ξ) they represent is given by

$$(\vec{x},\vec{\xi},\vec{x}\times\vec{\xi}) = \mathrm{R}_z(\phi)\cdot\mathrm{R}_x(\psi)\cdot\mathrm{R}_z(\theta). \tag{15}$$

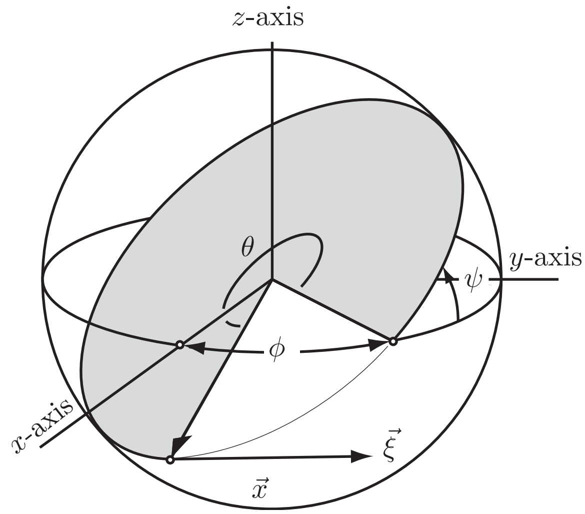

Figure 3: Euler angles ϕ, ψ and θ.

The map $(\vec{x}, \vec{\xi}) \mapsto (\theta, \psi, \phi)$ is a diffeomorphism between $\mathcal{U}$ and $(\mathbb{R}/2\pi\mathbb{Z}) \times (0, \pi) \times (\mathbb{R}/2\pi\mathbb{Z}) \cong \mathbb{T}^2 \times \mathbb{R}$.

Given this identification we can now define two numerical invariants of flat knots α relative to the equator ζ. By the lift of a unit speed parametrization, any flat knot $\alpha \in \Omega \setminus \Delta(\zeta)$ defines a closed curve $\hat{\alpha} : S^1 \to \mathcal{U}$. The numerical invariants are then the increments of the Euler angles θ and ϕ along $\hat{\alpha}$, which we will denote by $\Delta\theta(\alpha)$ and $\Delta\phi(\alpha)$, respectively. Both are integral multiples of 2π.

LEMMA 2.3. *If α is a satellite of ζ given by* (13) *then*

$$\pm\Delta\theta + \Delta\phi = 2q\pi, \tag{16a}$$

$$\Delta\theta = 2p\pi \tag{16b}$$

where $2p$ is the number of zeroes of $u \in C^2(\mathbb{R}/2q\pi\mathbb{Z})$. In the first equation one must take the "+ sign" if α has the same orientation as ζ, and the "− sign" otherwise.

Note that the number of zeroes of $u \in C^2(\mathbb{R}/2\pi\mathbb{Z})$ must always be even (assuming they are all simple zeroes, of course).

Proof. We project the sphere onto the cylinder $x^2 + y^2 = 1$ and write z and ϑ for the usual coordinates on this cylinder. We assume that α projects to the graph of $z = u(\vartheta)$ on the cylinder, and that u is a $2q\pi$ periodic function with simple zeroes only, and for which $|u(\vartheta)| + |u'(\vartheta)|$ is uniformly small. Let α have the same orientation as the equator (from west to east). We compute the Euler angles corresponding to the unit tangent vector to α at the point which projects to $(\vartheta_0, u(\vartheta_0))$ on the cylinder. In Figure 4 we have sketched the

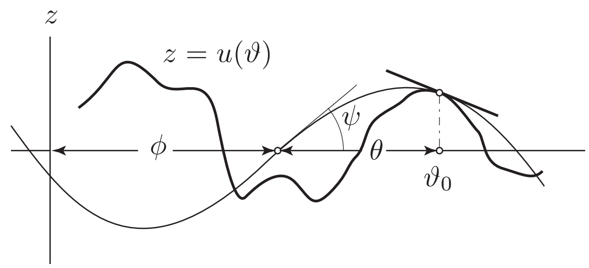

Figure 4: A great circle projected onto the cylinder.

great circle which passes through $(\vartheta_0, u(\vartheta_0))$ with slope $u'(\vartheta_0)$ as it appears in (ϑ, z) coordinates on the cylinder. Since great circles are intersections of planes through the origin with the sphere, they project to intersections of such planes with the cylinder, and are therefore graphs of $z = \psi \sin(\vartheta - \phi)$.

From Figure 4 one finds

$$\theta + \phi = \vartheta_0, \qquad u(\vartheta_0) = \psi \sin\theta, \qquad u'(\vartheta_0) = \psi\cos\theta, \tag{17}$$

so that

$$\theta = \arg(u'(\vartheta_0) + iu(\vartheta_0)). \tag{18}$$

From (17) we see that $\theta + \phi$ increases by $2q\pi$ along the curve α. To compute $\Delta\phi$ we use (18) and count the number of times the curve $u'(\vartheta_0) + iu(\vartheta_0)$ in the complex plane crosses the positive real axis. Every such crossing corresponds to a zero of u with positive derivative, and hence there are $\frac{2p}{2} = p$ of them. We conclude that $\Delta\theta = p \times 2\pi$, as claimed.

Similar arguments also allow one to find $\Delta\phi$ and $\Delta\theta$ if one gives α the orientation opposite to that of the equator. □

We have observed that $\mathcal{B}^+_{1,1}(\zeta)$ and $\mathcal{B}^-_{1,1}(\zeta)$ coincide. If p/q is any fraction in lowest terms then $\mathcal{B}^+_{p,q}(\zeta) = \mathcal{B}^-_{p,q}(\zeta)$ combined with (16a) implies $\Delta\theta = 0$, and hence $p = q$. Since $\gcd(p, q) = 1$ we conclude

LEMMA 2.4. *If ζ is a simple closed curve on S^2, and $\mathcal{B}^+_{p,q}(\zeta) = \mathcal{B}^-_{p,q}(\zeta)$ then $p = q = 1$.*

3. Curve shortening

3.1. *The gradient flow of the length functional.* Let g be a $C^{2,\mu}$ metric on the surface M. Then for any C^1 initial immersed curve γ_0 a maximal classical solution to curve shortening exists on a time interval $0 \le t < T(\gamma_0)$. We denote this solution by $\{\gamma_t : 0 \le t < T(\gamma_0)\}$. The solution depends continuously on the initial data $\gamma_0 \in \Omega$, so that curve shortening generates a continuous local semiflow

$$\Phi : \mathcal{D} \to \Omega, \qquad \Phi^t(\gamma_0) \stackrel{\text{def}}{=} \gamma_t,$$
$$\mathcal{D} = \{(\gamma, t) \in \Omega \times [0, \infty) \mid 0 \le t < T(\gamma)\}.$$

One can show that if $T(\gamma_0) < \infty$ then the geodesic curvature κ_{γ_t} of γ_t "blows-up" as $t \nearrow T(\gamma_0)$, i.e.

$$\lim_{t \nearrow T(\gamma_0)} \sup_{\gamma_t} |\kappa_{\gamma_t}| = \infty.$$

Since the geodesic curvature itself satisfies a parabolic equation

$$\frac{\partial \kappa_\gamma}{\partial t} = \frac{\partial^2 \kappa_\gamma}{\partial s^2} + \left(K \circ \gamma + \kappa_\gamma^2\right) \kappa_\gamma \tag{19}$$

($K \circ \gamma$ is the Gauss curvature of the surface evaluated along the curve) the maximum principle implies that one has the following lower estimate for the lifetime of any solution. If $T(\gamma_0) \le 1$ then

$$T(\gamma_0) \ge \frac{C}{\sqrt{\sup_{\gamma_t} |\kappa|}} \tag{20}$$

where C is some constant depending on $\sup_M |K|$ only. See [22] or [6].

The curve shortening flow on Ω provides a gradient flow for the length functional. Indeed, one has

$$\frac{dL(\gamma_t)}{dt} = -\int_{\gamma_t} (\kappa_{\gamma_t})^2 \, ds \tag{21}$$

where ds represents arclength along γ_t. Thus solutions of curve shortening do indeed always become shorter, unless γ_t is a geodesic, in which case the solution $\gamma_t \equiv \gamma_0$ is time independent. From the above description of $T(\gamma_0)$ one easily derives the following (see [23], [24], also [6], [7]).

LEMMA 3.1. *If* $T(\gamma_0) = \infty$ *then*

$$\lim_{t \to \infty} \sup_{\gamma_t} |\kappa_{\gamma_t}| = 0.$$

Moreover, any sequence $t_i \nearrow \infty$ *has a subsequence* t_i' *for which* $\gamma_{t_i'}$ *converges to some geodesic of* (M, g).

In other words, orbits of the curve shortening flow Φ which exist for all $t \ge 0$ have (compact) omega-limit sets in the sense of dynamical systems. Such ω-limit sets,

$$\omega(\gamma_0) \stackrel{\text{def}}{=} \{\gamma_* \in \Omega \mid \exists t_i \uparrow \infty : \gamma_{t_i} \to \gamma_*\}$$

are of course connected, and if the geodesics of (M, g) are isolated then any orbit of curve shortening either becomes singular or else converges to one geodesic.

The same is true for "ancient orbits," i.e. orbits $\{\gamma_t\}$ which are defined for all $t \le 0$ and for which $\sup_{t\le 0} L(\gamma_t) < \infty$. For such orbits one can define the α limit set

$$\alpha(\gamma_0) \stackrel{\text{def}}{=} \{\gamma_* \in \Omega \mid \exists t_i \searrow -\infty : \gamma_{t_i} \to \gamma_*\},$$

and this set consists of closed geodesics.

3.2. *Parabolic estimates.* Since curve shortening is a nonlinear heat equation solutions are generally smoother than their initial data. This provides a compactness property which we will use later to construct the Conley-index. There are various well-known ways of deriving the smoothing property of nonlinear heat equations. Here we show which estimate one can easily obtain assuming only that the metric g is C^2.

LEMMA 3.2. *If $\{\gamma_t \mid 0 \le t \le t_0\}$ is a solution of curve shortening whose curvature is bounded by $|\kappa| \le A$ at all times, then*

$$\int_{\gamma_t} \kappa_s^2 ds \le \frac{C}{t} \tag{22}$$

where the constant C only depends on A, t_0, the length L of $\gamma(0)$ and $\sup_M |K|$.

By adding a Nash-Moser iteration to the following arguments one could improve the estimate (22) to an L^∞ estimate for κ_s of the form $|\kappa_s| \le C/\sqrt{t}$. However, (22) will be good enough for us in this paper.

Proof. Let $\gamma : \mathbb{R}/\mathbb{Z}\times[0,T) \to M$ be a normal parametrization of a solution of curve shortening, i.e. one with $\partial_t\gamma \perp \partial_s\gamma$. Then the curvature κ satisfies (19), and using the commutation relation $[\partial_t, \partial_s] = \kappa^2\partial_s$ one obtains

$$\frac{\partial \kappa_s}{\partial t} = \frac{\partial^2 \kappa_s}{\partial s^2} + \frac{\partial}{\partial s}\big((K\circ\gamma)\kappa + \kappa^3\big). \tag{23}$$

The arclength ds on the curve evolves by $\frac{\partial}{\partial t} ds = -\kappa^2 ds$. Therefore we have

$$\begin{aligned} \frac{d}{dt}\int_{\gamma_t} (\kappa_s)^2 ds &= \int_{\gamma_t} \left(2\kappa_s\kappa_{st} - \kappa^2\kappa_s^2\right) ds \\ &= \int_{\gamma_t} \left(-2(\kappa_{ss})^2 + 5\kappa^2\kappa_s^2 - 2\kappa(K\circ\gamma)\kappa_{ss}\right) ds \\ &\le C + C\int_{\gamma_t} \kappa_s^2 ds - \int_{\gamma_t} (\kappa_{ss})^2 ds \end{aligned} \tag{24}$$

where the constant C only depends on A, L and $\sup_M |K|$.

By expanding $\kappa(\cdot, t)$ in a Fourier series in s one finds that

$$\left(\int_{\gamma_t} \kappa_s^2 ds\right)^2 \le \int_{\gamma_t} \kappa^2 ds \int_{\gamma_t} \kappa_{ss}^2 ds,$$

which implies

$$\int_{\gamma_t} \kappa_{ss}^2 ds \geq \frac{1}{C}\left(\int_{\gamma_t} \kappa_s^2 ds\right)^2$$

where the constant C only depends on $A = \sup|\kappa|$ and L. Combined with (24) this leads to a differential inequality for $\int \kappa_s^2 ds$,

$$\frac{d}{dt}\int_{\gamma_t} (\kappa_s)^2 ds \leq C + C\int_{\gamma_t} \kappa_s^2 ds - \frac{1}{C}\left(\int_{\gamma_t} \kappa_s^2 ds\right)^2.$$

Integration of this inequality gives (22). □

This lemma implies that for solutions with bounded curvature the curvature becomes Hölder continuous with exponent 1/2, since

$$\begin{aligned} |\kappa(P,t) - \kappa(Q,t)| &\leq \int_P^Q |\kappa_s| ds \\ &\leq \left(\int_P^Q \kappa_s^2 ds\right)^{1/2} \operatorname{dist}_{\gamma_t}(P,Q)^{1/2} \qquad \text{(Cauchy)} \\ &\leq \frac{C(L, A, \sup|K|)}{t} \operatorname{dist}_{\gamma_t}(P,Q)^{1/2}, \end{aligned} \tag{25}$$

$\operatorname{dist}_{\gamma_t}(P,Q)^{1/2}$ being the distance from P to Q along the curve γ_t.

3.3. *The nature of singularities in curve shortening.* Consider a solution $\{\gamma(t) : 0 \leq t < T\}$ of curve shortening with $T = T(\gamma_0) < \infty$. Then, as $t \nearrow T$, the curve γ_t converges to a piecewise smooth curve γ_T which has finitely many singular points $P_1, \ldots, P_m$; i.e. γ_T is the union of finitely many immersed arcs whose endpoints belong to $\{P_1, \ldots, P_m\}$.

Either γ_t shrinks to a point (in which case $m = 1$, and γ_T consists only of the point P_1), or else any neighborhood $\mathcal{U} \subset M^2$ of any of the P_i will contain a self-intersecting arc of γ_t for t sufficiently close to T. In other words, $\gamma_t \cap \mathcal{U}$ is the union of a finite number of arcs, at least one of which has a self-intersection (a parametrization $x \in \mathbb{R}/\mathbb{Z} \mapsto \gamma_t(x)$ of the curve will enter $\mathcal{U}$ and self-intersect before leaving the neighborhood).

This description of the singularities which a solution of curve shortening may develop follows from work of Grayson [23], [24]; see also [6], [7], [32] for a similar result applicable to more general flows; an alternative proof of the above result can now be given using the Hamilton-Huisken distinction between "type 1 and type 2" singularities (see [9] for a short survey), where we apply a monotonicity formula in the type 1 case, and either Hamilton's [25] or Huisken's isoperimetric ratios [26] in the type 2 case.

3.4. *Intersections and Sturm's theorem.* We recall Sturm's theorem [35] which states that if $u(x,t)$ is a classical solution of a linear parabolic equation

$$\frac{\partial u}{\partial t} = a(x,t)u_{xx} + b(x,t)u_x + c(x,t)u$$

on a rectangular domain $[x_0, x_1] \times [t_0, t_1]$, with boundary conditions

$$u(x_0,t) \neq 0,\ u(x_1,t) \neq 0, \quad \text{for } t_0 \leq t \leq t_1,$$

then the number of zeroes of $u(\cdot,t)$

$$z(u;t) \stackrel{\text{def}}{=} \#\{x \in [x_0,x_1] \mid u(x,t) = 0\}$$

is finite for any $t > t_0$, and does not increase as t increases. Moreover, at any moment t_* at which $u(\cdot,t_*)$ has a multiple zero, $z(u,t)$ drops. This theorem goes back to Sturm [35] who gave a rigorous proof assuming the solutions and coefficients are analytic functions, which has been rediscovered and reproved under weaker hypotheses many times since then. See [31], [29], [11].

In [10] we argue that Sturm's theorem may be considered as a "degenerate version" of the well-known principle that the local mapping degree of an analytic function $f : \mathbb{C} \to \mathbb{C}$ near any of its zeroes is always positive (so that one can count zeroes of f by computing winding numbers, etc.).

Using Sturm's theorem we proved the following in [6], [7].

LEMMA 3.3. *Any smooth solution $\{\gamma_t \mid 0 < t < T\}$ of curve shortening which is not a multiple cover of another solution, always has finitely many self-intersections, all of which are transverse, except at a discrete set of times $\{t_j\} \subset (0,T)$. At each time t_j the number of self-intersections of γ_t decreases.*

A similar statement applies to intersections of two different solutions: if $\{\gamma_t^1 \mid 0 < t < T\}$ and $\{\gamma_t^2 \mid 0 < t < T\}$ are solutions of curve shortening then they are transverse to each other, except at a discrete set of times $\{t_j\} \subset (0,T)$, and at each t_j the number of intersections of γ_t^1 and γ_t^2 decreases.

4. Curve shortening near a closed geodesic

4.1. *Eigenfunctions as (p,q) satellites.* Let $\gamma \in \Omega$ be a primitive closed geodesic of length L for a given $C^{2,\mu}$ metric g. We consider a C^1 neighborhood $\mathcal{U} \subset \Omega$ and parametrize it as in §2.1. Since the metric g is $C^{2,\mu}$, geodesics of g are $C^{3,\mu}$, and the unit normal to a geodesic will be $C^{2,\mu}$. We can therefore choose the local diffeomorphism $\sigma : \mathbb{R}/L\mathbb{Z} \times (-\delta, +\delta) \to M$ so that $x \mapsto \sigma(x,0)$ is a unit speed parametrization of γ and such that $\sigma_y(x,0)$ is a unit normal to γ at $\sigma(x,0)$.

The pullback of the metric under σ is

$$\sigma^*(g) = E(x,u)(dx)^2 + 2F(x,u)\,dx\,du + G(x,u)(du)^2,$$

for certain $C^{2,\mu}$ functions E, F, G.

One can map a C^1 neighborhood of $q \cdot \gamma$ in Ω onto a neighborhood of the origin in $C^2(\mathbb{R}/qL\mathbb{Z})$ via (6):

$$u \in C^2(\mathbb{R}/qL\mathbb{Z}) \mapsto \alpha_u \in \Omega, \qquad \alpha_u(x) = \sigma(x, u(x)). \tag{26}$$

In this chart the length functional $L : \Omega \to \mathbb{R}$ is given by

$$L(\alpha_u) = \int_0^{qL} \sqrt{E(x,u) + 2F(x,u)u_x + G(x,u)u_x^2}\, dx.$$

The curve α_u will be a geodesic if and only if u satisfies the Euler-Lagrange equations corresponding to L. Since we assume γ is already a geodesic, $u(x) \equiv 0$ satisfies the Euler-Lagrange equations. As is well-known, the second variation of L at $u = 0$ is then given by

$$d^2L(\gamma) \cdot (v,v) = \left.\frac{d^2L(\varepsilon v)}{d\varepsilon^2}\right|_{\varepsilon=0} = \int_0^{qL} \big(v'(x)^2 - K(\gamma(x))v(x)^2\big)\, dx$$

where $K(\gamma(x))$ is the Gauss curvature of (M, g) evaluated at $\gamma(x)$.

Consider the associated Hill's equation

$$\frac{d^2\varphi}{dx^2} + (Q(x) + \lambda)\varphi(x) = 0 \qquad (x \in \mathbb{R}) \tag{27}$$

where λ is an eigenvalue parameter, and where $Q(x) = K(\gamma(x))$ (although in what follows $Q \in C^0(\mathbb{R}/L\mathbb{Z})$ could be arbitrary).

Let $\varphi_i(x)$ be the solutions with initial conditions

$$\varphi_0(0) = 1, \quad \varphi_0'(0) = 0, \quad \varphi_1(0) = 0, \quad \varphi_1'(1) = 1, \tag{28}$$

and define the solution matrix

$$M(\lambda; x) = \begin{pmatrix} \varphi_0(x) & \varphi_1(x) \\ \varphi_0'(x) & \varphi_1'(x) \end{pmatrix}, \tag{29}$$

which belongs to $\mathrm{SL}(2, \mathbb{R})$.

If we identify the set of rays $\{\binom{ta}{tb} \mid t \geq 0, a^2 + b^2 = 1\}$ emanating from the origin in $\mathbb{R}^2$ with their intersections with the unit circle, then the linear transformation defined by $M(\lambda; x)$ also defines a homeomorphism of the unit circle to itself. This homeomorphism has a rotation number $\rho(\lambda, x)$, which is determined up to its integer part (see [18, §17.2]). To fix the integer part of $\rho(\lambda, x)$, we require that $\rho(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$ and that $\rho(\lambda, x)$ vary continuously with λ and x. The inverse rotation number of the geodesic mentioned in the introduction is precisely $\rho(\lambda = 0, x = L)$.

Since the coefficient $Q(x)$ is an L periodic function, one has

$$M(\lambda; qL) = M(\lambda; L)^q \tag{30}$$

and hence

$$\rho(\lambda, qL) = q\rho(\lambda, L). \tag{31}$$

The rotation number $\rho(\lambda, L)$ is a continuous nondecreasing function of the eigenvalue parameter λ, and thus for each fraction p/q the set of λ with $\rho(\lambda, L) = p/q$ is a closed interval $[\lambda^-_{p/q}, \lambda^+_{p/q}]$. Indeed, if $2p/q$ is not an integer, then $\lambda^-_{p/q} = \lambda^+_{p/q}$, and we just write $\lambda_{p/q}$.

The $\lambda^\pm_{p/q}$ depend on the potential Q, and depending on the context we will either write $\lambda_{p/q}(Q)$ or $\lambda_{p/q}(\gamma)$ if $Q = K \circ \gamma$ is the Gauss curvature evaluated along γ, as above.

Both for $\lambda = \lambda^-_{p/q}$, and $\lambda = \lambda^+_{p/q}$, Hill's equation (27) has a qL periodic solution which we denote by $\varphi^\pm_{p/q}(x)$. When $\lambda^-_{p/q} = \lambda^+_{p/q}$ both solutions $\varphi_i(\lambda; x)$ are qL periodic, and we let $\varphi^\pm_{p/q}(x)$ be φ_0, φ_1 respectively.

Let $E_{p/q}(Q)$ be the two dimensional subspace of $C^2(\mathbb{R}/qL\mathbb{Z})$ defined by

$$E_{p/q}(Q) \stackrel{\text{def}}{=} \left\{ c_+\varphi^+_{p/q}(x) + c_-\varphi^-_{p/q}(x) \;\middle|\; c_\pm \in \mathbb{R} \right\}. \tag{32}$$

This space is determined by $Q \in C^0(\mathbb{R}/L\mathbb{Z})$, i.e. does not require the geodesic γ or the surface M for its definition. It is the spectral subspace corresponding to the eigenvalues $\lambda^\pm_{p/q}$ of the unbounded operator $-\frac{d^2}{dx^2} - Q(x)$ in $L^2(\mathbb{R}/qL\mathbb{Z})$ and as such depends continuously on the potential $Q \in C^0(\mathbb{R}/qL\mathbb{Z})$.

Lemma 4.1. *Let α_ε be the satellite of γ given by $\alpha_{\varepsilon u}(x) = \sigma(x, \varepsilon u(x))$, with $u(x) \in E_{p/q}(K \circ \gamma)$, $u \neq 0$, and ε sufficiently small. Then $\alpha_{\varepsilon u}$ is a (p, q) satellite of γ, i.e. $\alpha_{\varepsilon u} \in \mathcal{B}_{p,q}(\gamma)$.*

Proof. The space $E_{p/q}(Q) \subset C^2(\mathbb{R}/qL\mathbb{Z})$ depends continuously on $Q \in C^0(\mathbb{R}/qL\mathbb{Z})$. For $Q(x) \equiv 0$ one has

$$E_{p/q}(0) = \{A \cos 2\pi \tfrac{p}{q} \tfrac{x}{L} + B \sin 2\pi \tfrac{p}{q} \tfrac{x}{L} \mid A, B \in \mathbb{R}\}.$$

Choose a continuous family of $\varphi_\theta \in E_{p/q}(\theta K \circ \gamma)$, $\varphi_\theta \neq 0$ with $\varphi_0(x) = \cos 2\pi \frac{p}{q} \frac{x}{L}$.

We must now show that for sufficiently small $\varepsilon \neq 0$ the corresponding curves

$$\alpha_{\varepsilon,\theta}(x) = \sigma(x, \varepsilon\varphi_\theta(x))$$

define flat knots relative to γ. To prove this we will show (i) that the graph of $\varphi_\theta(x)$ has no double zeroes (which implies that $\alpha_{\theta,\varepsilon}$ is never tangent to γ), and (ii) that the graphs of $\varphi_\theta(x)$ and $\varphi_\theta(x - kL)$ ($k = 1, 2, \ldots, q - 1$) have no tangencies (which implies that $\alpha_{\theta,\varepsilon}$ has no self-tangencies).

The following arguments are inspired by those in [12, §2].

If $\lambda^-_{p/q}(\theta K \circ \gamma) = \lambda^+_{p/q}(\theta K \circ \gamma)$, then φ_θ is a solution of Hill's equation (27) and cannot have a double zero without vanishing identically.

If $\lambda^-_{p/q}(\theta K \circ \gamma) \neq \lambda^+_{p/q}(\theta K \circ \gamma)$ then

$$\varphi_\theta(x) = c_-(\theta)\varphi^-_{p/q}(x) + c_+(\theta)\varphi^+_{p/q}(x)$$

for certain constants $c_\pm(\theta)$, at least one of which is nonzero. If one of these constants vanishes then φ_θ is again a solution of Hill's equation and therefore cannot have a double zero. If both coefficients $c_\pm$ are nonzero then we consider

$$u(t,x) = c_-(\theta)e^{\lambda^-_{p/q}t}\varphi^-_{p/q}(x) + c_+(\theta)e^{\lambda^+_{p/q}t}\varphi^+_{p/q}(x).$$

This function is a solution of the heat equation corresponding to Hill's equation, i.e.

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \theta K \circ \gamma(x) u,$$

and by Sturm's theorem the number of zeroes of $u(t,\cdot)$ must decrease at any moment t at which $u(t,\cdot)$ has a double zero. For $t \to \pm\infty$, $u(t,\cdot)$ is asymptotic to $c_\pm e^{\lambda^\pm t}\varphi^\pm_{p/q}(x)$, and since both $\varphi^\pm_{p/q}(x)$ have $2p$ zeroes in the interval $[0, qL)$ none of the intermediate functions $u(t,\cdot)$ can have a double zero. In particular $\varphi_\theta = u(0,\cdot)$ only has simple zeroes.

To prove (ii) one applies exactly the same arguments to the difference $\varphi_\theta(x) - \varphi_\theta(x - kL)$. The conclusion then is that this difference either only has simple zeroes (as desired), or else must vanish identically. To exclude the second possibility we observe that $\varphi_\theta(x) \equiv \varphi_\theta(x - kL)$ implies that φ_θ is an lL periodic function with $1 \le l < q$ some divisor of $\gcd(k,q)$. The number of zeroes of φ_θ then equals $\frac{q}{l}$ times the number of zeroes m of φ_θ in its minimal period interval $[0, lL)$. This number m is even, so the number of zeroes of φ_θ in the interval $[0, qL)$ is a multiple of $2q/l$. However, this number is $2p$ and so q/l must be a common divisor of p and q. This contradicts the hypothesis $\gcd(p,q) = 1$. □

4.2. *The linearized flow at a closed geodesic.* In the chart (26) curve shortening is equivalent to the following parabolic equation for $u(x,t)$ (see [6] and also §8.1):

$$u_t = \frac{u_{xx} + P(x,u) + Q(x,u)u_x + R(x,u)(u_x)^2 + S(x,u)(u_x)^3}{E(x,u) + 2F(x,u)u_x + G(x,u)(u_x)^2}. \tag{33}$$

The coefficients P, Q, R and S are C^1 functions of their arguments, and they satisfy

$$\begin{cases} P(x,0) = Q(x,0) = 0, \\ P_y(x,0) = K(\sigma(x,0)) \end{cases} \tag{34}$$

in which K is the Gauss curvature on the surface.

One can apply classical results on parabolic equations to deduce short-time existence for curve shortening from (33). In this section we shall use the local form of curve shortening to prove

Lemma 4.2. *If $\{\gamma_t \mid t \geq 0\}$ is an orbit of curve shortening which converges to a closed geodesic $\alpha \in \Omega$, then for t sufficiently large γ_t is a (p, q) satellite of α; i.e., $\gamma_t \in \mathcal{B}_{p,q}(\alpha)$ for some p, q. Moreover,*

$$\lambda^-_{p/q}(\alpha) \leq 0. \tag{35}$$

If $\{\gamma_t \mid t \leq 0\}$ is an "ancient orbit" of curve shortening with $\lim_{t\to-\infty} \gamma_t = \alpha$ for some closed geodesic $\alpha \in \Omega$, then for $-t$ sufficiently large γ_t is a (p, q) satellite of α for some p, q. In this case,

$$\lambda^+_{p/q}(\alpha) \geq 0. \tag{36}$$

Proof. We only prove the first statement; the second can be shown in the same way.

If γ_t converges to α in C^1 then we can choose coordinates as above, and for large t the curves γ_t correspond to a solution $u(x, t)$ of (33). This solution is defined for, say, $t \geq t_0$, and $u(\cdot, t) \to 0$ in $C^1(\mathbb{R}/\mathbb{Z})$ as $t \to \infty$. By parabolic estimates we also have $u(\cdot, t) \to 0$ in $C^2(\mathbb{R}/\mathbb{Z})$ as $t \to \infty$.

We can write (33) as

$$u_t = a(x, u, u_x)u_{xx} + b(x, u, u_x)u_x + c(x, u, u_x)u$$

where, using (34) and $E(x, 0) \equiv 1$,

$$\begin{aligned}
&a(x, u, p) = \left(E(x, u) + 2F(x, u)p + G(x, u)p^2\right)^{-1},\\
&b(x, 0, 0) = 0,\\
&c(x, 0, 0) = K(\sigma(x, 0)).
\end{aligned}$$

Thus (33) can be written as a quasilinear equation

$$u_t = \mathcal{A}(u)u$$

in which $\mathcal{A}(u)$ is the linear differential operator

$$\mathcal{A}(u) = a(x, u, u_x)\frac{d^2}{dx^2} + b(x, u, u_x)\frac{d}{dx} + c(x, u, u_x).$$

For $u = 0$ this operator reduces to

$$\mathcal{A}(0) = \frac{d^2}{dx^2} + K(\alpha(x))$$

whose spectrum we have just discussed.

Since u tends to zero, u asymptotically satisfies the equation $u_t = \mathcal{A}(0)u$, and thus for some $j \geq 0$ and some constant $C \neq 0$ one has

$$\lim_{t\to\infty} \frac{u(x, t)}{\|u(\cdot, t)\|_{L^2}} = C\varphi_j(x) \tag{37}$$

where $\varphi_j(x)$ is an eigenfunction of $\mathcal{A}(0)$ with $2j$ zeroes. See Lemmas 8.1 and 8.2. For large t the curve γ_t is therefore parametrized by

$$x \mapsto \sigma\big(x, \varepsilon(t)\{C\varphi_j(x) + o(1)\}\big),$$

where $\varepsilon(t) \to 0$ as $t \to \infty$. This implies that γ_t is a satellite of α.

If both eigenvalues $\lambda_\pm(p/q, K \circ \alpha)$ were positive then for large t one would have

$$\begin{aligned}\frac{d}{dt}\|u(\cdot,t)\|^2_{L^2} &= (u(t), \mathcal{A}(u(t))u(t))_{L^2} \\ &= (\lambda_\pm(p/q, K\circ\alpha) + o(1))\|u(\cdot,t)\|^2_{L^2} \\ &> 0\end{aligned}$$

which would keep $u(\cdot,t)$ from converging to zero. □

5. Loops

5.1. *Loops, simple loops, and filled loops.* Let $\gamma \in \Omega \setminus \Delta$ be a flat knot, and choose a parametrization $\gamma \in C^2(S^1, M)$, also denoted by γ. By definition a *loop* for γ is a nonempty interval $(a,b) \subset \mathbb{R}$ for which $\gamma(a) = \gamma(b)$ is a transverse self-intersection.

If we identify S^1 with $\partial\mathbb{D}$, where $\mathbb{D}$ is the unit disc in the complex plane, then $\gamma(a) = \gamma(b)$ implies that any simple loop $(a,b) \subset \mathbb{R}$ for γ defines a map $\bar\gamma : S^1 \to M$ via

$$\bar\gamma\left(e^{2\pi i \frac{t-a}{b-a}}\right) = \gamma(t), \quad \text{for } t \in (a,b).$$

By definition we will say that one can *fill in a loop* (a,b) if the map $\bar\gamma : \partial\mathbb{D} \to M$ can be extended to a local homeomorphism $\varphi : \mathbb{D} \to M$. We will always assume that a filling is at least C^1 on $\mathbb{D} \setminus \{1\}$, and that φ is a local diffeomorphism on $\mathbb{D} \setminus \{1\}$.

If $\bar\gamma : S^1 \to M$ is contractible, and one-to-one, then by the Jordan curve theorem one can fill $\bar\gamma$. We call such a loop an *embedded loop*.

Fillings come in two varieties which are distinguished by the way they approach the corner at the intersection $\gamma(a) = \gamma(b)$. The arcs $\gamma((a-\varepsilon, a+\varepsilon))$ and $\gamma((b-\varepsilon, b+\varepsilon))$ divide a small convex neighborhood of this intersection into four pieces ("quadrants"). The image $\varphi(D(1,\delta))$ of a small disk will intersect either one or three of these quadrants. If $\varphi(D(1,\delta))$ lies in one quadrant we call the corner convex, otherwise we call the corner concave.

5.2. *Continuation of loops and their fillings.* Let $\{\gamma_\theta \mid \theta \in [0,1]\} \subset \Omega \setminus \Delta$ be a smooth family of flat knots, and let γ_θ stand for smooth parametrizations of the corresponding curves. If $(a_0, b_0) \subset \mathbb{R}$ is a loop for γ_{θ_0} then, since all γ_θ have transverse self-intersections, the Implicit Function Theorem implies the

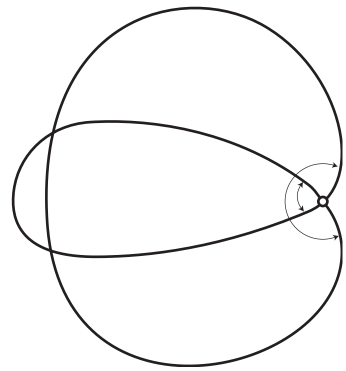

Figure 5: Convex and concave corners.

existence and uniqueness of smooth functions $a(\theta)$, $b(\theta)$ for which $(a(\theta), b(\theta))$ is a loop for $\gamma(\theta)$, and such that $a(\theta_0) = a_0$ and $b(\theta_0) = b_0$. Thus *any loop of a flat knot can be continued along homotopies of that flat knot.*

Now assume that the loop $(a_0, b_0) \subset \mathbb{R}$ of γ_{θ_0} has a filling: can one continue this filling in the same way? In general the answer is no, as the example in Figure 6 shows. It is also not true that embedded loops must remain embedded under continuation (see Figure 7)

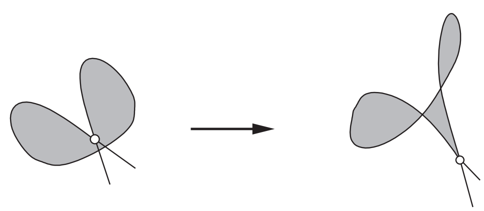

Figure 6: Inward corners may cut up fillings.

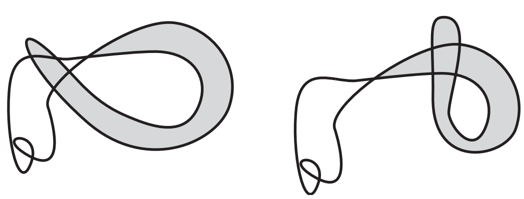

Figure 7: An embedded loop becomes nonembedded.

Lemma 5.1. *If the filling $\varphi_0 : \mathbb{D} \to M$ of the loop (a_0, b_0) has a convex corner, then there exists a continuous family of fillings $\varphi_\theta : \mathbb{D} \to M$ for the loops $(a(\theta), b(\theta))$ for all $\theta \in [0, 1]$.*

Proof. We may assume, by changing the parametrizations if necessary, that $a(\theta)$ and $b(\theta)$ are constant, so that (a,b) is a loop for each $\theta \in [0,1]$.

If one has a filling of a loop for some parameter value θ_0, then by constructing a tubular neighborhood of the arc $\gamma : [a,b] \to M$ one can adapt the given filling φ_0 to a filling φ_θ of the loops (a,b) for all θ in some interval $(\theta_0 - \varepsilon_0, \theta_0 + \varepsilon_0)$. To obtain a continuation from $\theta = 0$ all the way to $\theta = 1$ we must find a fixed lower bound for the size of the tubular neighborhoods. Such a lower bound then implies a lower bound for the length $2\varepsilon_0$ of the intervals on which one can construct local continuations, so that a finite number of such local continuations will take one from $\theta = 0$ to $\theta = 1$. We will therefore now describe the construction of the tubular neighborhoods of the γ_θ and the local continuations of the fillings in more detail.

Choose a suitable smooth metric g on the surface M. Then the Gauss curvature of (M,g) and geodesic curvatures of the γ_θ are uniformly bounded, say by some constant $\mathcal{K}$. We can therefore choose a small $\sigma > 0$ (much smaller than the injectivity radius of (M,g)) such that the intersection of any disk with radius σ at any point $P \in M$ with any of the curves γ_θ looks like a finite collection of straight line segments. More precisely, if we define the map $\phi_{P,\sigma}$ from the unit disc $\mathbb{D}_P \subset T_P M \cong \mathbb{R}^2$ to M, by $\phi_{P,\sigma}(x) = \exp_P(\sigma x)$, then the preimage $\phi_{P,\sigma}^{-1}(\gamma_\theta)$ is a finite collection of nearly straight arcs whose curvature is bounded by $C(\mathcal{K})\sigma$, which can be made arbitrarily small by decreasing σ.

For each $\theta \in [0,1]$ we construct a smooth vector field X_θ along γ_θ (i.e. $X_\theta : S^1 \to TM$ satisfies $X_\theta(t) \in T_{\gamma_\theta(t)}(M)$ for all $t \in S^1$), which is nowhere tangent to γ_θ, in particular $\angle(X_\theta(t), \gamma_\theta'(t)) \geq \delta$ for some constant $\delta > 0$. This δ can be chosen independently of θ. We can also choose the X_θ so that their derivatives are uniformly bounded, i.e. $|\nabla^j X_\theta| \leq C_j$ with C_j independent of θ. (Note that we do not assume that the X_θ vary continuously with θ.) Indeed, once one has constructed such a vector field for some value θ_1 of θ one can use the same vector field for all θ in an interval containing θ_1. A finite number of these intervals cover the interval $[0,1]$, so that we really only need a finite number of vector fields X_θ.

Let some $\theta_0 \in [0,1]$ be given, and let $\varphi_0 : \mathbb{D} \to M$ be a filling for the loop (a,b) of γ_{θ_0}. Since γ_{θ_0} is the image $\varphi_0(\mathbb{D})$ of the boundary of the unit disc one can define an "outward direction" at each $\gamma_{\theta_0}(t)$. We will assume that our vector field X along γ_{θ_0} is directed inward.

A tubular neighborhood is constructed from the mapping

$$S(t,s) = \exp_{\gamma_{\theta_0}(t)}(sX(t)).$$

This map is smooth from $S^1 \times \mathbb{R} \to M$. It is a local diffeomorphism on some neighborhood $\mathcal{U} = S^1 \times [-\rho, \rho]$ of $S^1 \times \{0\}$, where $\rho > 0$ is independent of θ_0.

If we choose X so that $X(a) = X(b)$, then this map is a local *homeomorphism* from the annulus $I \times [-\rho, \rho]$ to M, where $I = [a, b]/\{a, b\}$ (i.e. the interval $[a, b]$ with its endpoints identified so that $I \cong S^1$).

Now consider the curves $I \times \{s\}$ for $0 \le s < \rho$. For sufficiently small $s \ge 0$ there exist closed curves $\Gamma_s \subset \mathbb{D}$ for which

$$S(I \times \{s\}) = \varphi_{\theta_0}(\Gamma_s).$$

Each Γ_s is parametrized by $t \mapsto w(t, s)$, where w is the solution of

$$S(t, s) = \varphi_{\theta_0}(w(t, s)).$$

From $S(t, 0) = \gamma_{\theta_0}(t)$ it follows that $w = \exp\left(2\pi i \frac{t-a}{b-a}\right)$ is a solution for $s = 0$. Fix t and let s increase, starting at $s = 0$; then, since $\varphi_{\theta_0} : \mathbb{D} \to M$ is a local homeomorphism, one can continue the solution $w(t, s)$ to a solution $w = w(t, s) \in \operatorname{int}(\mathbb{D})$ for $0 \le s < \sigma(t) \le \rho$, where $\sigma(t)$ is a positive l.s.c. function of t. In particular, $\sigma(t)$ is bounded from below by some constant $\sigma > 0$. If for some $t \in I$ one has $\sigma(t) < \rho$, then as $s \uparrow \sigma(t)$ the solution $w(t, s)$ must tend to the boundary $\partial\mathbb{D}$ (otherwise one could continue the solution beyond $s = \sigma(t)$.)

It follows from $X(a) = X(b)$ that $w(a, s) \equiv w(b, s)$, and so $t \in [a, b] \mapsto w(t, s)$ parametrizes a closed curve Γ_s.

Proposition 5.2. *There exists a $\sigma' > 0$, independent of θ such that all Γ_s with $0 < s < \sigma'$ are disjoint embedded curves in $\mathbb{D}$.*

Proof. To begin, there is some $\sigma'' > 0$ such that none of the smooth immersed curves $t \in \mathbb{R}/\mathbb{Z} \mapsto S(t, s)$ with $|s| \le \sigma''$ has a self-tangency. This σ'' only depends on the choice of the vector fields X_θ, and we may thus assume that it is independent of θ.

The curves Γ_s are smooth, except at $w(a, s) = w(b, s)$, where they have a corner. Since the derivatives of the vector fields X_θ are bounded, we can find a $\sigma''' > 0$ independent of θ such that all curves $S(I \times \{s\})$ with $|s| \le \sigma'''$ have convex corners (in the sense that X_{θ_0} points "into the corner.") Hence the Γ_s also have convex corners for all $0 < s < \sigma'''$ for which they are defined.

Let

$$\sigma' = \min(\sigma, \sigma'', \sigma''').$$

As s increases from 0 to σ' the Γ_s must remain embedded, for the only way they can loose their embeddedness is by first forming a self-tangency. However, the smooth parts of the curves Γ_s are mapped to $S(I \times \{s\})$ which has no self-tangency. On the other hand the corner of Γ_s is convex, and so it cannot take part in a first self-tangency. Therefore the Γ_s remain embedded.

The Γ_s are nested. Indeed, they move with velocity

$$\frac{\partial w}{\partial s} = \mathrm{D}\varphi(w(t, s))^{-1} \frac{\partial S}{\partial s}$$

which is never tangent to Γ_s. Thus the Γ_s always move in the same direction, which must be inward, since they start at $\Gamma_0 = \partial\mathbb{D}$.

Being nested, the Γ_s can never reach the boundary $\partial\mathbb{D}$ again, and hence they exist for all $s \in (0, \sigma')$. □

Conclusion of proof of Lemma 5.1. By "straightening" the curves Γ_s, we see that the above construction allows us to modify the filling φ_0 so that on the annulus $e^{-\sigma'/2} \le |w| \le 1$ it is given by

$$\varphi_0(re^{i\phi}) = S(a + \frac{\phi}{2\pi}(b-a), -\ln r). \tag{38}$$

For this φ_0 the curves Γ_s are circles centered at the origin. Then we use this same expression (38) to extend φ_0 to a local homeomorphism $\bar{\varphi}_0 : \mathbb{D}_{e^{\sigma'}} \to M$.

Since all $\gamma|_{[a,b]}$ with θ close to θ_0 are transverse to the vector field X, the preimage under $\bar{\varphi}_0$ of a nearby loop $\gamma_\theta\big|_{[a,b]}$ appears as a graph $r = r(\phi)$ in polar coordinates. One easily adapts the filling φ_0 to a filling of $\gamma_\theta\big|_{[a,b]}$ by first mapping the unit disk $\mathbb{D}$ to the region enclosed by the polar graph $r = r(\phi)$, and then composing with $\bar{\varphi}_0$. The length of the interval of θ's for which one can do this is bounded from below by some $\delta > 0$ which is independent of θ, and hence a finite number of these local continuations will allow one to fill $\gamma_\theta\big|_{[a,b]}$ for all $\theta \in [0, 1]$. □

5.3. *Loops and singularities in curve shortening.* In §3.3 we considered a solution $\{\gamma_t \mid 0 \le t < T\}$ of curve shortening which becomes singular at time $t = T$ without shrinking to a point. In the notation of §3.3 we recalled that Grayson's work implies that for every neighborhood $\mathcal{U}$ of a singular point P_j there is a time $T_{\mathcal{U}} \in (0, T)$ such that for $T_{\mathcal{U}} < t < T$ the curve γ_t has a loop $(a', b') \subset [0, 1)$ with $\gamma_t([a', b'])$ contained in $\mathcal{U}$. Such a loop need not be simple, but one can easily extract a subloop $(a, b) \subset (a', b')$ which is simple. Since $\gamma_t|(a, b)$ is simple it is also a fillable loop. Still, the loop could have a nonconvex corner, but if this is the case, and if the neighborhood $\mathcal{U}$ is homeomorphic to a disc, then we claim one can find another loop, which is contained in $\mathcal{U}$, which is simple, and whose filling has a convex corner.

Indeed, let $\mathcal{R} \subset M$ be the region enclosed by the loop, and let A be the (nonconvex) corner point of $\mathcal{R}$. Since A is a nonconvex corner point the two arcs of $\gamma \setminus \partial\mathcal{R}$ enter *into* the region $\mathcal{R}$ (see Figure 8). There are now two possibilities:

Case 1. If one of these arcs exits $\mathcal{R}$ again (say, at $B \in \partial\mathcal{R}$) without first forming a self-intersection, then the arc AB divides $\mathcal{R}$ into two pieces, the boundary of one of which is a simple loop with a convex corner B.

Case 2. If both arcs starting at A self-intersect before leaving $\mathcal{R}$, then each of these arcs contains a simple loop whose area is strictly smaller than that

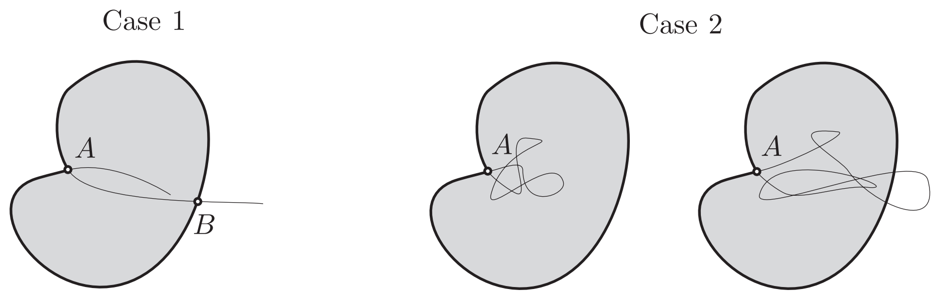

Figure 8: Finding a fillable loop with a convex corner.

of $\mathcal{R}$. If this smaller loop still does not have a convex corner then we repeat the argument, thereby obtaining a nested sequence of smaller simple loops. Since γ only has finitely many loops this sequence must terminate either with a simple loop with a convex corner, or with a loop as in Case 1.

Thus we can refine the description of singularities in §3.3 to the following:

LEMMA 5.3. *If* $\{\gamma_t \mid 0 \le t < T\}$ *is a solution of curve shortening which becomes singular at* $t = T$*, then for any* $\varepsilon > 0$ *there exists a* $t_\varepsilon \in (0, T)$ *such that* γ_{t_ε} *has a convexly fillable loop with area no more than* ε*.*

5.4. *Decrease of area of small loops.* Let $\{\gamma_t \mid t_0 \le t < t_1\}$ be a solution of curve shortening with $\gamma \in \Omega \setminus \Delta$ for all $t \in (t_0, t_1)$. Assume γ_{t_0} has a fillable loop with a convex corner. Then one can continue this loop for all values of $t \in (t_0, t_1)$. Let $\varphi_t : \mathbb{D}^2 \to M$ be a filling of these loops. Since φ_t is a local diffeomorphism away from $1 \in \mathbb{D}^2$, we can pull the metric back from M to $\mathbb{D}^2$ and define the area form dS_t and Gauss curvature K_t of $\varphi_t^*(g)$, as well as the geodesic curvature κ_t and arc length ds_t of the boundary $\partial\mathbb{D}^2$. The Gauss-Bonnet formula states that

$$\int_{\partial\mathbb{D}} \kappa_t ds_t + \iint_{\mathbb{D}} K_t dS_t + \theta_{\mathrm{ext}} = 2\pi.$$

Here θ_{ext} is the exterior angle at the corner of the filling. See Figure 9.

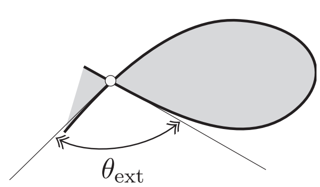

Figure 9: Definition of θ_{ext}

Using this we find that the area $A(t)$ of the filling φ_t satisfies

$$\frac{dA(t)}{dt} = \frac{d}{dt} \iint_{\mathbb{D}^2} \varphi_t^*(dS) = -\int_{\partial\mathbb{D}^2} \kappa ds = -2\pi + \theta_{\mathrm{ext}} + \int\int K dS.$$

Since $0 < \theta_{\mathrm{ext}} < \pi$ this implies

$$\frac{dA(t)}{dt} < -\pi + (\sup_M K)A(t).$$

Define

$$\varepsilon(g) = \frac{\pi}{2\sup_M K}$$

if $\sup_M K > 0$ and $\varepsilon(g) = \infty$ otherwise. We may then conclude:

LEMMA 5.4. *Let $\gamma_0 \in \Omega \setminus \Delta$ have a convexly filled loop with area at most $\varepsilon(g)$, and consider the corresponding solution $\{\gamma_t \mid 0 \le t < T\}$ of curve shortening. As long as the solution stays in $\Omega \setminus \Delta$ one can continue the loop, and its area satisfies*

$$A'(t) \le \frac{\pi}{2}, \text{ and } A(t) \le \varepsilon(g) - \frac{\pi}{2}t. \tag{39}$$

In particular the solution must either become singular or cross Δ before $t_* = \frac{2\varepsilon(g)}{\pi}$.

6. Definition of the Conley index of a flat knot

6.1. *The boundary of a relative flat knot type.* Let $\mathcal{B} \subset \Omega \setminus \Delta(\Gamma)$ be a relative flat knot type, for some $\Gamma = \{\gamma_1, \ldots, \gamma_N\} \subset \Omega$. Throughout we will make the following assumption concerning multiple covers

(40) If $\alpha = m \cdot \beta \in \Omega$, $m \ge 2$, $\beta \in \Omega$, is tranverse to all γ_i then $\alpha \notin \bar{\mathcal{B}}$.

We mention some examples.

6.1.1. (p, q) *satellites.* If M is the sphere and ζ is the equator, then consider $\mathcal{B} = \mathcal{B}_{p,q}(\zeta)$. Let $\mathcal{U}$ be the subset of the unit tangent bundle which consists of all vectors not tangent to ζ. We have seen in §2.8 that $\mathcal{U}$ has the homotopy type of $\mathbb{T}^2$. Any $\alpha \in \Omega$ which is transverse to ζ lifts to a curve $\hat{\alpha}$ in $\mathcal{U}$, and hence defines a homotopy class $[\hat{\alpha}]$ in $\pi_1(\mathbb{T}^2) \cong \mathbb{Z}^2$. The homotopy class $[\hat{\alpha}]$ does not depend on $\alpha \in \mathcal{B}$, and hence on $\alpha \in \bar{\mathcal{B}}$. Since $\gcd(p, q) = 1$ this homotopy class is not a multiple of any other element of $\pi_1(\mathbb{T}^2)$, and therefore α cannot be a multiple of another curve. We conclude that the relative flat knot types $\mathcal{B}_{p,q}(\zeta)$ satisfy condition (40).

This example is easily generalized to any relative flat knot type $\mathcal{B} \subset \Omega \setminus \Delta(\Gamma)$. Define $\mathcal{U}$ to be the unit tangent bundle of M with the tangent vectors to the $\pm\gamma_i$ removed, and assume that the homotopy class $[\hat{\alpha}] \in \pi_1(\mathcal{U})$ is not a multiple of any other element of $\pi_1(\mathcal{U})$. Then $\bar{\mathcal{B}}$ cannot contain multiple covers transverse to the $\gamma_i \in \Gamma$.

6.1.2. *Simple closed curves.* Let $\mathcal{S}$ be the set of simple closed curves on $M = S^2$. If $\alpha = m\beta$ is a multiple cover, then any $\alpha' \in \Omega$ near α must have at least one self-intersection. Hence $\mathcal{S}$ satisfies condition (40). However in this case $\mathcal{U}$ is the entire unit tangent bundle $T^1S^2 \cong \mathbb{RP}^3$ whose fundamental group is $\mathbb{Z}_2$, in which $2 \cdot 1 = 0$ and $3 \cdot 1 = 1$, i.e. in which all elements are nontrivial multiples. So just like $\mathcal{B}_{p,q}(\zeta)$ the flat knot type $\mathcal{S}$ satisfies (40), but it does so for different reasons.

6.1.3. *Free satellites.* Let p, q with $\gcd(p, q) = 1$ be given and consider the set $\mathcal{B}$ of all $\alpha \in \Omega$ for which a simple closed curve $\zeta \in \Omega$ exists such that α is a (p, q) satellite of ζ. Since any two simple closed curves can be deformed into each other by isotopy of S^2, the set $\mathcal{B}$ is a connected component of $\Omega \setminus \Gamma$, and hence the set of curves which are (p, q) satellites is a flat knot type. Note that, in contrast with the example from §6.1.1 the curve ζ here is not fixed, and the set Γ is empty. Our current set $\mathcal{B}$ is a flat knot type, while the set $\mathcal{B}$ from §6.1.1 was only a relative flat knot type.

For any simple closed curve ζ the q fold cover $q \cdot \zeta$ lies on the boundary $\partial\mathcal{B}$ since one can approximate it by (p, q) satellites of ζ. Since there are no γ_i in this example, this flat knot type does not satisfy the condition (40).

We consider the closure $\bar{\mathcal{B}}$ of $\mathcal{B}$ in Ω and define

$$\begin{aligned}\hat{\mathcal{B}} &= \bar{\mathcal{B}} \setminus \{\pm m \cdot \gamma_i \mid m \geq 2, i = 1, \ldots, N\},\\ \partial\hat{\mathcal{B}} &= \hat{\mathcal{B}} \cap \partial\mathcal{B}.\end{aligned}$$

LEMMA 6.1. *Let $g \in \mathcal{M}_\Gamma$. For any $\alpha \in \partial\hat{\mathcal{B}}$ a $t_\alpha > 0$ exists such that $\Phi^{(0,t_\alpha)}(\alpha) \subset \mathcal{B} \cup \big(\Omega \setminus \bar{\mathcal{B}}\big)$.*

Recall that the curve shortening flow Φ^t was defined in 3.1.

Proof. Since $\alpha \in \bar{\mathcal{B}}$ the curve α has only finitely many crossings with any of the γ_i. Hence for some $t_1 > 0$ all $\Phi^t(\alpha)$ with $0 < t < t_1$ are transverse to all γ_i. If α is not primitive, then condition (40) implies that $\Phi^t(\alpha) \in \Omega \setminus \bar{\mathcal{B}}$. If α is primitive, then we may assume that the $\Phi^t(\alpha)$ with $0 < t < t_1$ also have transverse self-intersections. Hence $\Phi^t(\alpha) \in \Omega \setminus \Delta \subset \mathcal{B} \cup \big(\Omega \setminus \bar{\mathcal{B}}\big)$. □

The following lemma states that orbit segments cannot touch the boundary of a flat knot type $\mathcal{B}$ "from the inside".

LEMMA 6.2. *Let $g \in \mathcal{M}_\Gamma$. If $\Phi^{[0,t]}(\alpha) \subset \hat{\mathcal{B}}$ and $t > 0$ then $\Phi^s(\alpha) \in \mathcal{B}$ for all $s \in (0, t)$.*

Proof. Suppose for some $s \in (0, t)$ one has $\Phi^s(\alpha) \in \partial\hat{\mathcal{B}}$. Then $\Phi^s(\alpha)$ cannot be a multiple cover by condition (40). By the Sturmian theorem $\Phi^{s'}(\alpha)$ is a flat knot rel Γ for all $s' \neq s$ close to s, and either the number of self-intersections or the number of intersections of $\Phi^{s'}(\alpha)$ with some $\gamma_i \in \Gamma$ must drop as s' crosses s. This contradicts $\Phi^{[0,t]}(\alpha) \subset \hat{\mathcal{B}}$. □

We define the *exit set* of $\hat{\mathcal{B}}$ to be the set $\mathcal{B}^-$ consisting of those $\alpha \in \partial\hat{\mathcal{B}}$ for which $\Phi^{(0,t_\alpha)}(\alpha) \subset \Omega \setminus \bar{\mathcal{B}}$. The complement $\mathcal{B}^+ = \partial\hat{\mathcal{B}} \setminus \mathcal{B}^-$ is called the *entry set*.

LEMMA 6.3. *The sets $\mathcal{B}^\pm$ do not depend on the metric $g \in \mathcal{M}_\Gamma$ chosen in their definition.*

LEMMA 6.4. *$\mathcal{B}^-$ is a closed subset of $\hat{\mathcal{B}}$.*

We prove these lemmas in reverse order.

6.1.4. *Proof of Lemma* 6.4. We first show that $\mathcal{B}^+$ is open in $\partial\hat{\mathcal{B}}$. Let $\alpha \in \mathcal{B}^+$ be given. Then $\Phi^{(0,t_\alpha)} \subset \mathcal{B}$ and in particular $\Phi^{t_\alpha/2}(\alpha) \in \mathcal{B}$. By continuity of the local semiflow Φ there is an open neighborhood $\mathcal{N} \subset \Omega$ containing α such that $\Phi^{t_\alpha/2}(\mathcal{N}) \subset \mathcal{B}$. Suppose some $\alpha' \in \mathcal{N}$ belongs to $\mathcal{B}^-$. Then there is a small $t' \in (0, t_\alpha/2)$ such that $\Phi^{t'}(\alpha') \in \Omega \setminus \bar{\mathcal{B}}$. By continuity of Φ again, there is an $\alpha'' \in \mathcal{N} \cap \mathcal{B}$ with $\Phi^{t'}(\alpha'') \in \Omega \setminus \bar{\mathcal{B}}$. But then $\alpha'' \in \mathcal{B}$, $\Phi^{t'}(\alpha'') \in \Omega \setminus \bar{\mathcal{B}}$, and $\Phi^{t_\alpha/2}(\alpha'') \in \mathcal{B}$. This contradicts the Sturmian theorem. $\square$

6.1.5. *Proof of Lemma* 6.3. We classify the possible curves $\alpha \in \partial\hat{\mathcal{B}}$ as follows.

(1) α is primitive and transverse to all γ_i, but α has a self-tangency.

(2) α is primitive and tangent to some γ_i (but $\alpha \neq \pm\gamma_i$ by condition (40))

(3) $\alpha = m \cdot \beta$ ($m \geq 2$, $\beta \in \Omega$) is a multiple cover. In this case α must be tangent to at least one of the γ_i, by condition (40).

The curves in Case 3 all belong to $\mathcal{B}^-$, for under the curve shortening flow they remain multiple covers, while they instantaneously become transverse to the γ_i, so that condition (40) forces them to leave $\bar{\mathcal{B}}$.

The following proposition shows that we have in Case 1,

$$\alpha \in \mathcal{B}^- \Leftrightarrow \mathrm{Cross}(\alpha, \alpha) < m_0,$$

while in Case 2 we have

$$\alpha \in \mathcal{B}^- \Leftrightarrow \exists i : \mathrm{Cross}(\alpha, \gamma_i) < m_i.$$

Thus we have a description of the exit set which is independent of the chosen metric g. $\square$

PROPOSITION 6.5. *Let $\alpha \in \hat{\mathcal{B}}$ be primitive with $\alpha \neq \pm\gamma_i$ for any i. If $\mathrm{Cross}(\alpha, \alpha) = m_0(\mathcal{B})$ and $\mathrm{Cross}(\alpha, \gamma_i) = m_i(\mathcal{B})$ for $i = 1, \ldots, N$, then for some $\varepsilon > 0$ one has $\Phi^{(0,\varepsilon)}(\alpha) \subset \mathcal{B}$.*

Proof. Since $\Phi^{(0,\varepsilon]}(\alpha) \subset \Omega \setminus \Delta(\gamma_1, \dots, \gamma_N)$ we either have $\Phi^{(0,\varepsilon]}(\alpha) \subset \mathcal{B}$ or $\Phi^{(0,\varepsilon]}(\alpha) \subset \Omega \setminus \bar{\mathcal{B}}$. We must show the latter cannot hold. Suppose it does hold. Then let $\alpha_n \in \mathcal{B}$ be a sequence with $\alpha_n \to \alpha$. Since $\Phi^\varepsilon(\alpha) \in \Omega \setminus \bar{\mathcal{B}}$ one also has $\Phi^\varepsilon(\alpha_n) \in \Omega \setminus \bar{\mathcal{B}}$ for large enough $n \in \mathbb{N}$. Thus the orbit $\Phi^t(\alpha_n)$ crosses $\Delta(\gamma_1, \dots, \gamma_N)$ for some $t \in [0, \varepsilon]$. By the Sturmian theorem one then has

$$\#\big(\Phi^\varepsilon(\alpha_n) \cap \Phi^\varepsilon(\alpha_n)\big) < m_0 \text{ or } \exists i : \#\big(\Phi^\varepsilon(\alpha_n) \cap \gamma_i\big) < m_i.$$

On the other hand $\Phi^\varepsilon(\alpha) \in \Omega \setminus \Delta(\gamma_1, \dots, \gamma_N)$ so that for sufficiently large $n \in \mathbb{N}$ one has

$$\begin{aligned} \text{either} \quad & \#\big(\Phi^\varepsilon(\alpha_n) \cap \Phi^\varepsilon(\alpha_n)\big) = \#\big(\Phi^\varepsilon(\alpha) \cap \Phi^\varepsilon(\alpha)\big) < m_0 \\ \text{or} \quad & \exists i : \#\big(\Phi^\varepsilon(\alpha_n) \cap \gamma_i\big) = \#\big(\Phi^\varepsilon(\alpha) \cap \gamma_i\big) < m_i. \end{aligned}$$

If we now let $\varepsilon \downarrow 0$ then we get

$$\mathrm{Cross}(\alpha, \alpha) < m_0 \text{ or } \exists i : \mathrm{Cross}(\alpha, \gamma_i) < m_i.$$

Thus we have a contradiction, and the proposition is proved. □

Given a metric g we now define

$$\mathcal{B}^\ell(g, \varepsilon) = \left\{ \alpha \in \mathcal{B} \;\middle|\; \begin{array}{l} \alpha \text{ has a filled loop with} \\ \text{a convex corner, and area} \le \varepsilon \end{array} \right\}, \tag{41}$$

$$\hat{\mathcal{B}}^\ell(g, \varepsilon) = \text{the closure of } \mathcal{B}^\ell(g, \varepsilon) \text{ in } \Omega \setminus \{\pm q\gamma_i \mid q \in \mathbb{N}, 1 \le i \le N\}, \tag{42}$$

and we call

$$h(\mathcal{B}) = \left[\hat{\mathcal{B}} \Big/ \hat{\mathcal{B}}^\ell(g, \varepsilon) \cup \mathcal{B}^- \right],$$

the Conley-index of the component $\mathcal{B}$. Here for any closed subset A of a topological space X, $[X/A]$ stands for the homotopy type of the pointed quotient space X/A. See [17].

Lemma 6.6. *The set $\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g, \varepsilon)$ is positively invariant relative to $\hat{\mathcal{B}}$.*

Proof. Let $\Phi^{[0,t]}(\alpha) \subset \hat{\mathcal{B}}$ with $t > 0$ and $\alpha \in \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g, \varepsilon)$ be given.

By Lemma 6.2 we have $\Phi^s(\alpha) \in \mathcal{B}$ for $0 < s < t$.

Fix an $s \in (0, t)$ and choose a sequence $\alpha_n \to \alpha$ with $\alpha_n \in \mathcal{B}^\ell(g, \varepsilon)$. Since $\Phi^s(\alpha) \in \mathcal{B}$ continuity of the semiflow implies $\Phi^s(\alpha_n) \in \mathcal{B}$ for large enough n. Foward invariance of $\mathcal{B}^\ell(g, \varepsilon)$ in $\mathcal{B}$ then implies $\Phi^s(\alpha_n) \in \mathcal{B}^\ell(g, \varepsilon)$. Taking the limit as $n \to \infty$ one finds $\Phi^s(\alpha) \in \hat{\mathcal{B}}^\ell(g, \varepsilon)$ for any $s \in (0, t)$. Taking another limit $s \to t$ one finds that $\Phi^t(\alpha) \in \hat{\mathcal{B}}^\ell(g, \varepsilon)$. □

Lemma 6.7. *The Conley index $h(\mathcal{B})$ does not depend on the metric $g \in \mathcal{M}_\Gamma$ or the choice of $\varepsilon > 0$, as long as $\varepsilon < \varepsilon(g)$.*

This lemma justifies the absence of g and ε in our notation "$h(\mathcal{B})$" for the index.

The proof of this lemma is essentially found in [17]. Our observation here is that although we do not have the desired local compactness[1] we are only trying to prove that the index is independent of the "index pair" for a small class of index pairs.

We split the proof of Lemma 6.7 into two pieces. It will be convenient to write

$$\mathcal{H}(g,\varepsilon) = \hat{\mathcal{B}} \Big/ \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)$$

so that we have defined the Conley index of $\mathcal{B}$ to be the homotopy type of the quotient $\mathcal{H}(g,\varepsilon)$, and we must now show that this homotopy type does not depend on $g \in \mathcal{M}_\Gamma$ or $\varepsilon \in (0, \varepsilon(g))$.

6.1.6. *$h(\mathcal{B})$ does not depend on ε.* Let $0 < \varepsilon_1 < \varepsilon_2 < \varepsilon(g)$ be given. Then trivially we have the inclusion $\mathcal{B}^\ell(g,\varepsilon_1) \subset \mathcal{B}^\ell(g,\varepsilon_2)$ which leads to a natural mapping

$$\mathcal{H}(g,\varepsilon_1) \xrightarrow{f} \mathcal{H}(g,\varepsilon_2).$$

(Whenever $A_1 \subset A_2 \subset X$ are closed subsets there is a natural mapping $X/A_1 \to X/A_2$.)

We will show that this mapping is a homotopy equivalence. For every $\gamma \in \hat{\mathcal{B}}$ we define

$$T_*(\gamma) = \inf\{t \geq 0 \mid \Phi_t(\gamma) \in \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon_1)\} \tag{43}$$

with the understanding that $T_*(\gamma) = \infty$ if $\Phi_t(\gamma)$ never reaches the exit set or $\hat{\mathcal{B}}^\ell(g,\varepsilon_1)$.

PROPOSITION 6.8. *The function $T_* : \hat{\mathcal{B}} \to [0,\infty]$ is continuous.*

Proof. We check that both conditions $T_*(\gamma) < M$ and $T_*(\gamma) > M$ define open subsets of $\hat{\mathcal{B}}$.

If $T_*(\gamma) < M$ for some $0 < M < \infty$, then $\Phi^{T_*(\gamma)}(\gamma)$ belongs to $\mathcal{B}^-$ or $\hat{\mathcal{B}}^\ell(g,\varepsilon_1)$. In the first case the orbit immediately leaves $\mathcal{B}$, and so there exists a $t_0 \in (T_*(\gamma), M)$ with $\Phi^{t_0}(\gamma) \in \Omega \setminus \bar{\mathcal{B}}$. By continuity of the semiflow Φ the same is then true for all γ' near γ, so that $T_*(\gamma') < t_0 < M$ holds on a neighborhood of γ.

Consider the second case, in which $\Phi^{T_*(\gamma)}(\gamma) \in \hat{\mathcal{B}}^\ell(g,\varepsilon_1)$. If $T_*(\gamma) = 0$ then it is possible that $\gamma = \Phi^{T_*(\gamma)}(\gamma)$ lies on $\mathcal{B}^+$. When this happens $\Phi^t(\gamma)$

[1]K. Rybakowski has developed a version of Conley's theory for local semiflows on complete metric spaces, but we were unable to verify his "admissibility condition," mainly because $\hat{\mathcal{B}}$ can contain arbitrarily long curves, and, possibly, geodesics of arbitrary length.

must immediately enter $\mathcal{B}$ and hence$\mathcal{B}^\ell(g,\varepsilon_1)$, by forward invariance of $\hat{\mathcal{B}}^\ell(g,\varepsilon)$ relative to $\hat{\mathcal{B}}$. If on the other hand $T_*(\gamma) > 0$ then $\Phi^{T_*(\gamma)}(\gamma)$ cannot lie on $\mathcal{B}^+$. By assumption it does not lie on $\mathcal{B}^-$ either, and thus it lies in $\mathcal{B}^\ell(g,\varepsilon_1)$. It follows from Lemma 5.4 that at $t = T_*(\gamma)$ the orbit $\Phi^t(\gamma)$ develops a convexly filled loop with area $\le \varepsilon_1$, and that for $t > T_*(\gamma)$ the loop has area $\le \varepsilon_1 - \frac{\pi}{2}(t - T_*(\gamma))$ which is strictly less than ε_1. Invoking continuity of the semiflow we conclude again that this condition also holds for γ' near γ.

Conversely, if $T_*(\gamma) > M$, then the (compact) orbit segment $\{\gamma_t \mid 0 \le t \le M\}$ is contained in $\hat{\mathcal{B}} \setminus \hat{\mathcal{B}}^\ell(g,\varepsilon_1)$ which is open relative to $\hat{\mathcal{B}}$. Once more continuity of the semiflow guarantees that this is also the case for γ' close to γ. □

It follows from Lemma 5.4 that for all $\gamma \in \mathcal{B}^\ell(g,\varepsilon_2)$ one has $T_*(\gamma) \le \frac{2}{\pi}(\varepsilon_2 - \varepsilon_1)$. By continuity this also holds for all $\gamma \in \hat{\mathcal{B}}^\ell(g,\varepsilon_2)$. Now define

$$T_0(\gamma) = \min\left(\frac{2}{\pi}(\varepsilon_2 - \varepsilon_1), T_*(\gamma)\right)$$

and consider the following homotopy ($0 \le \lambda \le 1$):

$$G_\lambda : \mathcal{B}^- \cup \hat{\mathcal{B}} \to \mathcal{B}^- \cup \hat{\mathcal{B}}, \qquad \gamma \mapsto \Phi^{\lambda T_0(\gamma)}(\gamma).$$

Then

- G_0 is the identity map on $\mathcal{B}^- \cup \hat{\mathcal{B}}$,
- G_λ maps $\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)$ to itself for every $\varepsilon \in (0, \varepsilon(g))$ (by forward invariance of $\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell$, Lemma 6.6),
- G_1 maps $\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon_1)$ into $\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon_2)$

and it is easily verified from these facts that G_1 is a homotopy inverse of f.

6.1.7. *$h(\mathcal{B})$ does not depend on the metric.* Let $g_1, g_2 \in \mathcal{M}_\Gamma$ be two given metrics. Then, since the surface M is compact there exists a constant $A > 0$ such that one has $g_1 \le Ag_2$ and $g_2 \le Ag_1$ pointwise on M. In particular the area form of either metric is bounded by A^2 times the area form of the other. We therefore have the following inclusions

$$\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g_1,\varepsilon) \subset \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g_2,A^2\varepsilon) \subset \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g_1,A^4\varepsilon) \subset \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g_2,A^6\varepsilon)$$

and corresponding natural maps

$$\mathcal{H}(g_1,\varepsilon) \xrightarrow{f} \mathcal{H}(g_2,A^2\varepsilon) \xrightarrow{g} \mathcal{H}(g_1,A^4\varepsilon) \xrightarrow{h} \mathcal{H}(g_2,A^6\varepsilon).$$

Now it follows from the previous section that for sufficiently small $\varepsilon > 0$ the compositions $g \circ f$ and $h \circ g$ are homotopy equivalences, so that g has a left and right homotopy inverse. Hence g is a homotopy equivalence. □

6.2. *Virtual satellites and the modified Conley index of a relative flat knot.* Let $\mathcal{B} \subset \Omega \backslash \Delta(\Gamma)$ be a relative flat knot for some $\Gamma = \{\gamma_1, \dots, \gamma_N\} \subset \Omega$. Define $\mathcal{M}_\Gamma$ as in §1.3, and, as in §1.3, order the γ_i so that for $i = 1, \dots, m$ there exists p_i/q_i with

$$q_i \gamma_i \in \partial \mathcal{B} \text{ and } \mathcal{B} \subset \mathcal{B}_{p_i,q_i}(\gamma_i) \tag{44}$$

while no such p_i/q_i exist for $i = m+1, \dots, n$. By Lemma 2.2 the p_i/q_i are uniquely determined.

We impose the nonresonance condition (3) and for any $I \subset \{1, \dots, m\}$ we define $\mathcal{M}(\alpha; I)$ for $\alpha \in \mathcal{B}$ as in §1.3. Since $\mathcal{M}(\alpha; I)$ does not depend on $\alpha \in \mathcal{B}$ we will write $\mathcal{M}(\mathcal{B}; I)$ for $\mathcal{M}(\alpha; I)$. Our discussion of the rotation number in §4.1 shows that condition (4) is equivalent to

$$\lambda^+_{p_i/q_i}(\gamma_i) < 0 \text{ for } i \in I, \text{ and } \lambda^-_{p_i/q_i}(\gamma_i) > 0 \text{ for } i \in I^c, \tag{45}$$

where $I^c = \{1, \dots, m\} \setminus I$.

For the moment write $q = q_i$ and $\gamma = \gamma_i$. Let $\mathcal{U} \subset \bar{\mathcal{B}}$ be a closed neighborhood in $\bar{\mathcal{B}}$ of $q\gamma$ which is small enough for $q\gamma$ to be the only geodesic in $\mathcal{U}$, and for $\mathcal{U} \cap \hat{\mathcal{B}}^\ell(g, \varepsilon)$ to be empty (for some $\varepsilon \in (0, \varepsilon(g))$ which we keep fixed throughout this section).

Define

$$\mathcal{U}_{\text{transient}} = \{\alpha \in \mathcal{U} \cap \bar{\mathcal{B}} \mid \exists t > 0 : \Phi^t(\alpha) \in \mathcal{B} \setminus \mathcal{U}\} \tag{46}$$

$$\mathcal{U}^\# = \mathcal{U} \setminus (\mathcal{U}_{\text{transient}} \cup \{q\gamma\}) \,. \tag{47}$$

The conditions $\alpha \in \bar{\mathcal{B}}$ and $\Phi^t(\alpha) \in \mathcal{B}$ imply that $\Phi^{(0,t]} \subset \mathcal{B}$, since orbits cannot leave and then enter $\bar{\mathcal{B}}$ again. Thus $\mathcal{U}^\#$ consists of those $\alpha \in \mathcal{U}$ which do not leave $\mathcal{U}$ before leaving $\mathcal{B}$.

Clearly $\mathcal{U}_{\text{transient}}$ is open, so that $\mathcal{U}^\#$ is closed in $\hat{\mathcal{B}}$ and $\mathcal{U}^\# \cup \{q\gamma\}$ is closed in $\bar{\mathcal{B}}$.

By construction $\mathcal{U}^\#$ is positively invariant relative to $\hat{\mathcal{B}}$; $\mathcal{U} \setminus \{q \cdot \gamma\}$ is positively invariant relative to $\hat{\mathcal{B}}$ if and only if $\mathcal{U} = \mathcal{U}^\# \cup \{q \cdot \gamma\}$, or, equivalently,

$$\mathcal{U}^\# = \mathcal{U} \setminus \{q \cdot \gamma\}.$$

LEMMA 6.9. *$\mathcal{U}^\# \cup \{q\gamma\}$ is a neighborhood in $\bar{\mathcal{B}}$ of $q \cdot \gamma$.*

Proof. If $\mathcal{U}^\# \cup \{q\gamma\}$ is not a neighborhood of $q\gamma$ then a sequence $\alpha_n \in \bar{\mathcal{B}} \backslash \mathcal{U}^\#$ with $\lim_{n\to\infty} \alpha_n = q\gamma$ must exist. Since $\mathcal{U}$ is assumed to be a neighborhood we may assume that all $\alpha_n \in \mathcal{U}$, and thus $\alpha_n \in \mathcal{U}_{\text{transient}}$. Then $t'_n > 0$ exist with $\Phi^{t'_n}(\alpha_n) \in \mathcal{B} \setminus \mathcal{U}$. Choose t_n to be the largest $t \in (0, t'_n)$ with $\Phi^{[0,t_n]}(\alpha_n) \subset \mathcal{U}$. In particular $\Phi^{t_n}(\alpha_n) \in \partial \mathcal{U}$.

Since $q\gamma$ is a fixed point for curve shortening, we have $\lim_{n\to\infty} t_n = \infty$. By parabolic estimates 3.2 we can extract a convergent subsequence of the sequence of solutions $\{\beta_n(t) \stackrel{\text{def}}{=} \Phi^{t_n+t}(\alpha_n) \mid -t_n < t \leq 0\}$. The limit is an "ancient orbit" $\{\beta(t) \mid -\infty < t \le 0\}$ which remains in $\mathcal{U}$, and which reaches $\partial\mathcal{U}$ at $t = 0$. The α-limit of such an orbit must be $q\gamma$ (being the only closed geodesic in $\mathcal{U}$) but this contradicts $\lambda^+_{p/q}(\gamma) < 0$ and Lemma 4.2. $\square$

For any neighborhood $\mathcal{U} \subset \bar{\mathcal{B}}$ of $q\gamma$ and $T > 0$ we define

$$\mathcal{U}^{\#T} = \{\Phi^T(\alpha) \mid \Phi^{[0,T]}(\alpha) \subset \mathcal{U}^{\#}\}.$$

As T increases the set $\mathcal{U}^{\#T}$ shrinks. In general $\mathcal{U}^{\#T}$ is not a neighborhood of $q \cdot \gamma$; in fact, due to the regularizing effect of the heat flow, $\mathcal{U}^{\#T}$ will have empty interior.

LEMMA 6.10. *Let $\mathcal{U}, \mathcal{V} \subset \bar{\mathcal{B}}$ be neighborhoods of $q\gamma$. Then for sufficiently large $T > 0$ one has $\mathcal{U}^{\#T} \subset \mathcal{V}^{\#}$ and $\mathcal{V}^{\#T} \subset \mathcal{U}^{\#}$.*

Proof. We need only prove the first inclusion, and we may of course assume that the neighborhoods $\mathcal{U} = \mathcal{U}^{\#} \cup \{q\gamma\}$, $\mathcal{V} = \mathcal{V}^{\#} \cup \{q\gamma\}$ are positively invariant relative to $\mathcal{B}$.

Arguing by contradiction we assume that there exists a sequence $\alpha_k \in \mathcal{U}$ with $\Phi^{[0,k]}(\alpha_k) \subset \mathcal{U} \setminus \mathcal{V}$. Parabolic estimates yield an *a priori* bound for $\frac{\partial k}{\partial s}$ on the curves $\Phi^1(\alpha_k)$, and thus we can extract a convergent subsequence from the solutions $\beta_k(t) = \Phi^{t+1}(\alpha_k)$ of curve shortening. The limit would then be an orbit of curve shortening which stays in $\overline{\mathcal{U} \setminus \mathcal{V}}$, in particular its ω-limit would be a closed geodesic other than $q\gamma$ in $\mathcal{U}$, which by assumption does not exist. $\square$

Let $I \subset \{1, \ldots, m\}$ and $g \in \mathcal{M}_\Gamma(\mathcal{B}; I)$ be given. For each $i \in I$ we choose a sufficiently small neighborhood $\mathcal{U}_i$ of $q_i\gamma_i$ and we set

$$\mathcal{U}^I \stackrel{\text{def}}{=} \bigcup_{i \in I} \mathcal{U}_i^{\#}.$$

We will assume that the $\mathcal{U}_i \setminus \{q_i\gamma_i\}$ are forward invariant relative to $\hat{\mathcal{B}}$, i.e. $\mathcal{U}_i = \mathcal{U}_i^{\#}$.

Definition 6.11. The *modified Conley index* of the relative flat knot type $\mathcal{B}$ is

$$h^I(\mathcal{B}) = \left[\hat{\mathcal{B}} \Big/ \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g, \varepsilon) \cup \mathcal{U}^I \right].$$

Our previously defined Conley index $h(\mathcal{B})$ is contained in this definition as the special case in which $I \subset \{1, \ldots, m\}$ is empty.

LEMMA 6.12. *For sufficiently small $\mathcal{U}_i$ and $\varepsilon > 0$ the index $h^I(\mathcal{B})$ does not depend on either ε, the metric $g \in \mathcal{M}_\Gamma(\mathcal{B}; I)$ or the neighborhoods $\mathcal{U}_i$.*

Proof. We may assume that $\mathcal{U} \supset \mathcal{V}$ for otherwise we choose a smaller neighborhood $\mathcal{W} \subset \mathcal{U} \cap \mathcal{V}$ and compare the indices $h^I(\mathcal{B})$ obtained by using $\mathcal{U}$ and $\mathcal{V}$ with the index obtained by using $\mathcal{W}$.

Choose a sufficiently large $T > 0$ so that $\mathcal{U}^{\#T} \subset \mathcal{V}$ and as before in (43) define

$$T_*(\alpha) = \inf\{t \geq 0 \mid \Phi^t(\alpha) \in \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell\}.$$

We showed in Proposition 6.8 that the exit time $T_*(\alpha)$ is a continuous function with values in $[0, \infty]$. The family of maps

$$F_\theta(\alpha) \stackrel{\text{def}}{=} \Phi^{\theta \min(T_*(\alpha), T)}(\alpha)$$

with $\theta \in [0,1]$ is therefore a continuous homotopy $F_\theta : \mathrm{id} \cong F_1$ of maps of the pairs $(\hat{\mathcal{B}}, \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$ and $(\hat{\mathcal{B}}, \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$. From $\mathcal{U}^{\#T} \subset \mathcal{V}$ we conclude that F_1 maps the quotient $\hat{\mathcal{B}}/(\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$ to $\hat{\mathcal{B}}/(\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{V})$, and is a homotopy inverse for the inclusion induced map from $\hat{\mathcal{B}}/(\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{V})$ to $\hat{\mathcal{B}}/(\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$. □

6.3. *The Conley index of a virtual satellite.*

LEMMA 6.13. *The homotopy type of* $\mathcal{U}_i^{\#}/(\mathcal{U}_i^{\#} \cap \mathcal{B}^-)$ *is that of* $S^1 \times S^{2p_i-1}/S^1 \times \{\mathrm{pt}\}$. *Consequently the homotopy type of* $\mathcal{U}^I/(\mathcal{U}^I \cap \mathcal{B}^-)$ *is given by*

$$\bigvee_{i=1}^{m} [S^1 \times S^{2p_i-1}/S^1 \times \{\mathrm{pt}\}].$$

We will call the homotopy type of $\mathcal{U}_i^{\#}/(\mathcal{U}_i^{\#} \cap \mathcal{B}^-)$ the Conley index of the virtual satellite of $q_i\gamma_i$ in $\mathcal{B}$.

In the following proof we omit the subscript i and write $\mathcal{U}$ instead of $\mathcal{U}_i$, etc.

The same arguments as in Corollary 6.12 show that $\mathcal{U}^{\#}/(\mathcal{U}^{\#} \cap \mathcal{B}^-)$ is independent of both the metric g, provided γ is a geodesic with $\lambda^+_{p/q}(\gamma) < 0$, and the neighborhood $\mathcal{U}$, provided it is sufficiently small, meaning that it should not contain any other closed geodesics besides $q\gamma$ and be disjoint from $\hat{\mathcal{B}}^\ell$. Thus we may choose our metric so that a neighborhood of γ in the surface M is isometric to a part of the surface of revolution whose metric is given by

$$(ds)^2 = e^{y^2/2}\{(dx)^2 + (dy)^2\}, \quad (x, y) \in (\mathbb{R}/\mathbb{Z}) \times \mathbb{R},$$

where the "waist" $y = 0$ corresponds to γ. Curves $\alpha \in \Omega$ which are C^1 close to $q\gamma$ are then given by graphs of functions $u \in C^2(\mathbb{R}/q\mathbb{Z})$, $\alpha(x) = (x, u(x))$ (such a graph wraps itself q times around the waist $\{y = 0\}$). In this section we will identify a small neighborhood of $q \cdot \gamma \in \Omega$ with an open neighborhood of $u \equiv 0$ in $C^2(\mathbb{R}/q\mathbb{Z})$ without explicitely mentioning the identification again.

Curve shortening for such graphs is equivalent to the PDE

$$\frac{\partial u}{\partial t} = \frac{u_{xx}}{1+u_x^2} - u. \tag{48}$$

We choose

$$\mathcal{U} = \mathcal{N}_\sigma \cap \bar{\mathcal{B}} \tag{49}$$

where

$$\mathcal{N}_\sigma = \{u \in C^2(\mathbb{R}/q\mathbb{Z}) \mid \sup_x |u(x)| \le \sigma, \sup_x |u'(x)| \le \sigma\}, \tag{50}$$

and σ is sufficiently small.

LEMMA 6.14. *$\mathcal{N}_\sigma$ is invariant for the curve shortening flow, so that $\mathcal{U}$ contains no transient part, i.e. $\mathcal{U} = \mathcal{U}^\# \cup \{q\gamma\}$.*

Proof. The maximum principle implies that any solution of (48) with $|u(x,0)| \le \sigma$ satisfies $|u(x,t)| \le \sigma e^{-t}$, since $\pm\sigma e^{-t}$ are sub- and supersolutions for (48).

By differentiating (48) one finds that $v = u_x$ satisfies

$$v_t = \frac{v_{xx}}{1+u_x^2} - \frac{2u_x}{(1+u_x^2)^2} v_x - v$$

so that the maximum principle again implies that $\sup_x |v(x,0)| \le \sigma$ leads to $\sup_x |v(x,t)| \le \sigma e^{-t}$. □

We will identify $\Delta \subset \Omega$ with those $u \in C^2(\mathbb{R}/q\mathbb{Z})$ which correspond to a curve $\alpha_u \in \Delta$.

LEMMA 6.15. *Both $\mathcal{B}^-$ and Δ are cones in $C^2(\mathbb{R}/q\mathbb{Z})$.*

Proof. A function $u \in C^2(\mathbb{R}/q\mathbb{Z})$ belongs to Δ if it either has a multiple zero or if for some $k = 1, \ldots, q-1$ the function $u(x) - u(x-k)$ has a multiple zero. This clearly holds for u if and only if it holds for λu, for any $\lambda \ne 0$. Thus Δ is a cone.

Near $q\gamma$ the set $\mathcal{B}^-$ consists of those $u \in \Delta$ which have fewer self-intersections, or fewer intersections with $u = 0$ than a general $u \in \mathcal{B}$ has. This condition also holds for both u and λu or for neither. □

Any $u \in C^2(\mathbb{R}/q\mathbb{Z})$ has a Fourier series of the form

$$u(x) = \sum_{k=0}^{\infty} \mathfrak{Re}\left(u_k e^{2k\pi i x/q}\right) \tag{51}$$

with $u_0 \in \mathbb{R}$ and $u_k \in \mathbb{C}$ for $k \ge 2$. The embedding $C^2 \hookrightarrow W^{2,2}$ implies that

$$\int_0^q \{u(x)^2 + \frac{q^4}{16\pi^4} u''(x)^2\} dx = \sum_{k=0}^{\infty} (1+k^4)|u_k|^2 < \infty. \tag{52}$$

We now define

$$\mathcal{S}_\varepsilon = \left\{ u \in C^2(\mathbb{R}/q\mathbb{Z}) \;\middle|\; \sum_{k=0}^{\infty}(1+k^4)|u_k|^2 = \varepsilon^2 \right\};$$

i.e., $\mathcal{S}_\varepsilon$ is the intersection with $C^2(\mathbb{R}/q\mathbb{Z})$ of the sphere of radius ε in $W^{2,2}(\mathbb{R}/q\mathbb{Z})$ with norm given by (52).

Since $W^{2,2} \hookrightarrow C^1$ one has $\mathcal{S}_\varepsilon \subset \mathcal{N}_\sigma$ for small enough $\varepsilon > 0$.

Lemma 6.15 implies that $(\mathcal{S}_\varepsilon \cap \hat{\mathcal{B}}, \mathcal{S}_\varepsilon \cap \mathcal{B}^-)$ is a deformation retract of $(\mathcal{U}^\#, \mathcal{U}^\# \cap \mathcal{B}^-)$, the deformation going along rays through the origin in $C^2(\mathbb{R}/q\mathbb{Z})$. We therefore have a homotopy equivalence

$$\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-) \cong (\mathcal{S}_\varepsilon \cap \mathcal{U}^\#)/(\mathcal{S}_\varepsilon \cap \mathcal{U}^\# \cap \mathcal{B}^-).$$

For small enough $\varepsilon > 0$ one has $\mathcal{S}_\varepsilon \subset \mathcal{N}_\sigma$, so that $\mathcal{S}_\varepsilon \cap \mathcal{U}^\# = \mathcal{S}_\varepsilon \cap \hat{\mathcal{B}}$, and $\mathcal{S}_\varepsilon \cap \mathcal{U}^\# \cap \mathcal{B}^- = \mathcal{S}_\varepsilon \cap \mathcal{B}^-$. Hence we have a further homotopy equivalence

$$\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-) \cong (\mathcal{S}_\varepsilon \cap \hat{\mathcal{B}})/(\mathcal{S}_\varepsilon \cap \mathcal{B}^-).$$

The linear heat equation induces a continuous semiflow on $\mathcal{S}_\varepsilon$: for any $u \in \mathcal{S}_\varepsilon$ let

$$u(t,x) = \sum_{k=0}^{\infty} \mathfrak{Re}\left(u_k e^{2k\pi i x/q - 4k^2\pi^2/q^2 t}\right) \tag{53}$$

be the solution of $u_t = u_{xx}$ starting from u, and define $\Psi^t(u)$ to be the radial projection of $u(t,\cdot)$ onto $\mathcal{S}_\varepsilon$, so that

$$\big(\Psi^t u\big)(x) = \varepsilon \frac{u(t,x)}{\|u(t,\cdot)\|_{W^{2,2}}}.$$

We will refer to Ψ^t as the projected heat flow.

The essential insight which allows us to determine the homotopy type of $(\mathcal{S}_\varepsilon \cap \hat{\mathcal{B}})/(\mathcal{S}_\varepsilon \cap \mathcal{B}^-)$ and hence of $\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-)$ is that $(\mathcal{S}_\varepsilon \cap \hat{\mathcal{B}}, \mathcal{S}_\varepsilon \cap \mathcal{B}^-)$ turns out to be an index pair for the projected heat flow $\Psi^t : \mathcal{S}_\varepsilon \to \mathcal{S}_\varepsilon$ which isolates the invariant set

$$\mathcal{C} \stackrel{\text{def}}{=} \left\{ \mathfrak{Re}\left(u_p e^{2\pi i p x/q}\right) \;\middle|\; u_p \in \mathbb{C}, |u_p| = \frac{\varepsilon}{\sqrt{1+p^4}} \right\}.$$

This invariant set is a normally hyperbolic circle whose unstable manifold is $2p$ dimensional, so one expects its Conley index to be $\big[S^1 \times S^{2p-1}/S^1 \times \mathrm{pt}\big]$. Since we do not have the required compactness hypothesis of [17], we must prove these statements by hand, essentially verifying that Conley's arguments still go through in our setting.

For $u \in \mathcal{S}_\varepsilon$ we define the quantities

$$w_+(u) = \varepsilon^{-2} \sum_{k>p} \left(1+k^4\right) |u_k|^2,$$
$$w_-(u) = \varepsilon^{-2} \sum_{0\le k<p} \left(1+k^4\right) |u_k|^2,$$
$$w_p(u) = \varepsilon^{-2} \left(1+p^4\right) |u_p|^2.$$

By definition we have

$$w_-(u) + w_p(u) + w_+(u) = 1$$

for all $u \in \mathcal{S}_\varepsilon$.

LEMMA 6.16. *Along any orbit $\Psi^t(u)$ of the projected heat flow one has*

$$\frac{dw_+(\Psi^t u)}{dt} \le -Cw_+(1-w_+) < 0, \tag{54}$$

$$\frac{dw_-(\Psi^t u)}{dt} \ge -Cw_-(1-w_-) > 0. \tag{55}$$

Proof. Let $u(t,x)$ be the solution to the linear heat equation starting at $u \in \mathcal{S}_\varepsilon$ given by (53), so that $u(t,x) = \sum_{k\ge 0} \mathfrak{Re}(u_k(t)e^{2\pi ikx/q})$ with $u_k(t) = e^{-4\pi^2k^2/q^2 t}u_k(0)$. Then one has

$$\begin{aligned}\frac{dw_+(u(t,\cdot))}{dt} &= \sum_{k>p} \left(1+k^4\right) \mathfrak{Re}(2u_k'(t)u_k(t)) \\ &\le -\frac{8\pi^2(p+1)^2}{q^2} w_+(u(t,\cdot))\end{aligned}$$

and

$$\begin{aligned}\frac{dw_-(u(t,\cdot))}{dt} &\ge -\frac{8\pi^2(p-1)^2}{q^2} w_-(u(t,\cdot)), \\ \frac{dw_p(u(t,\cdot))}{dt} &= -\frac{8\pi^2p^2}{q^2} w_p(u(t,\cdot)).\end{aligned}$$

Using

$$w_\pm(\Psi^t(u)) = \frac{w_\pm(u(t,\cdot))}{w_+(u(t,\cdot)) + w_p(u(t,\cdot)) + w_-(u(t,\cdot))}$$

one then arrives at (54) and (55) with $C = 8\pi^2(2p \pm 1)/q^2$. □

Consider the sets

$$\mathcal{V}_\rho = \{u \in \mathcal{S}_\varepsilon \cap \hat{\mathcal{B}} \mid w_+(u) \le \rho\},$$
$$\mathcal{V}_\rho^- = \{u \in \mathcal{V}_\rho \mid w_-(u) \ge \rho\}.$$

The differential inequalities (54), (55) imply that $(\mathcal{V}_\rho, \mathcal{V}_\rho^-)$ is an index pair. It isolates the same invariant set $\mathcal{C}$ as $(\mathcal{U}^\#, \mathcal{U}^\# \cap \mathcal{B}^-)$, so one expects $\mathcal{V}_\rho/\mathcal{V}_\rho^-$ and $\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-)$ to have the same homotopy type. To prove this we exhibit homotopy equivalences obtained by "flowing along" with Ψ^t

$$\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}_-) \longrightarrow \mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-), \tag{56}$$

and

$$\mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-) \longrightarrow \mathcal{V}_\rho/\mathcal{V}_\rho^-. \tag{57}$$

6.3.1. *Construction of the homotopy equivalence* (56). We define

$$t_*(u) = \inf\{t \geq 0 \mid w_+(\Psi^t(u)) \leq \rho\}.$$

PROPOSITION 6.17. *The function t_* is continuous and finite on $\mathcal{S}_\varepsilon \cap \hat{\mathcal{B}}$.*

Proof. We first observe that one has $w_+(u) < 1$ for every $u \in \mathcal{U}^\#$. Indeed, if $w_+(u) = 1$ then $w_-(u) = w_p(u) = 0$ and so the Fourier series (51) only contains terms with $k \geq p+1$. Then u has at least $2(p+1)$ sign changes and cannot belong to $\hat{\mathcal{B}}$ or $\mathcal{U}^\# \subset \hat{\mathcal{B}}$.

The differential inequality (54) implies that $w_+(\Psi^t(u))$ will decrease to ρ in finite time so that t_* is finite. Moreover $\frac{d}{dt}w_+ < 0$ implies that the time at which $w_+(\Psi^t(u))$ becomes equal to ρ depends continuously on u. □

The family of maps $G_\theta(u) = \Psi^{\theta t_*(u)}(u)$ is a continuous homotopy $G_\theta : \mathrm{id} \cong G_1$. The final map G_1 sends $\mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-)$ to $\mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-)$ and is a homotopy inverse for the inclusion induced map $\mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-) \to \mathcal{U}^\#/(\mathcal{U}^\# \cap \mathcal{B}^-)$.

6.3.2. *Construction of the homotopy equivalence* (57). To construct a map from left to right in (57) we observe

PROPOSITION 6.18. *If $\rho > 0$ is small enough then $\mathcal{V}_\rho \cap \mathcal{B}^- \subset \mathcal{V}^-$.*

Proof. We consider the sets

$$\mathcal{W}_\rho = \{u \in \mathcal{S}_\varepsilon \mid w_+(u) \leq \rho, w_-(u) \leq \rho\}, \tag{58}$$

$$\mathcal{W}_\rho^- = \{u \in \mathcal{S}_\varepsilon \mid w_+(u) \leq \rho, w_-(u) = \rho\}. \tag{59}$$

By definition $\mathcal{W}_\rho$ is a $W^{2,2}$ neighborhood of Γ which can be made as small in $W^{2,2}$ as desired by decreasing ρ. Since Γ is compact, and since Δ is closed in C^1 and thus also in $W^{2,2}$, we conclude that, for sufficiently small $\rho > 0$, $\mathcal{W}_\rho$ and Δ are disjoint. Since $\mathcal{V}_\rho \setminus \mathcal{V}_\rho^- \subset \mathcal{W}_\rho$ and $\mathcal{B}^- \subset \Delta$ the proposition follows. □

For small enough ρ the proposition guarantees that we have an inclusion induced map $\mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-) \longrightarrow \mathcal{V}_\rho/\mathcal{V}_\rho^-$. A homotopy inverse for this map can again be found by following the flow. Define an "exit time"

$$t_*(u) \stackrel{\text{def}}{=} \inf\{t \geq 0 \mid \Psi^t(u) \in \mathcal{B}^-\}.$$

If we allow $t_*(u) = \infty$ in case the orbit $\Psi^t(u)$ never hits $\mathcal{B}^-$ then $t_*(u)$ depends continuously on $u \in \mathcal{U}$, again because orbits cross $\mathcal{B}^-$ in a topologically transverse way (this is the same argument as in Proposition 6.8).

PROPOSITION 6.19. *If $w_-(u) > 0$ then $t_*(u) < \infty$.*

Proof. Let k_0 be the smallest integer with $u_{k_0} \neq 0$. Then $\Psi^t(u) = \varepsilon \frac{u(t,\cdot)}{\|u(t,\cdot)\|_{2,2}}$ with $u(t,x)$ given by (53). For large t the dominant term in (53) is the term with $k = k_0$, so that

$$\lim_{t\to\infty} \Psi^t(u) = \text{Const} \cdot \mathfrak{Re}\left(u_{k_0} e^{2\pi i k_0 x/q}\right).$$

Since $w_-(u) > 0$ we have $k_0 < p$, and hence for large t, $\Psi^t(u)$ has less than $2p$ sign changes so that $\Psi^t(u)$ cannot lie in $\bar{\mathcal{B}}$ anymore. The only way $\Psi^t(u)$ can leave $\mathcal{B}$ is by crossing $\mathcal{B}^-$ first. □

We can now define the following family of maps,

$$G_\theta(u) = \Psi^{\theta\eta(w_-(u))t_*(u)}(u)$$

in which $\eta : \mathbb{R} \to \mathbb{R}$ is a continuous nondecreasing function with $\eta(w) \equiv 0$ for $w \leq \rho/2$ and $\eta(w) \equiv 1$ for $w \geq \rho$. Thus $\eta(w_-(u))$ vanishes in the region $w_-(u) \leq \rho/2$ while $t_*(u)$ is continuous for $w_-(u) > 0$ so that the product $\eta(w_-(u))t_*(u)$ is continuous everywhere.

The G_θ are maps of the pairs $(\mathcal{V}_\rho, \mathcal{V}_\rho^-)$ and $(\mathcal{V}_\rho, \mathcal{V}_\rho \cap \mathcal{B}^-)$ respectively, and the final map G_1 sends $(\mathcal{V}_\rho, \mathcal{V}_\rho^-)$ to $(\mathcal{V}_\rho, \mathcal{V}_\rho \cap \mathcal{B}^-)$. It therefore provides a homotopy inverse for the inclusion induced map $\mathcal{V}_\rho/(\mathcal{V}_\rho \cap \mathcal{B}^-) \to \mathcal{V}_\rho/\mathcal{V}_\rho^-$.

6.3.3. *Computation of the homotopy type of $\mathcal{V}_\rho/\mathcal{V}_\rho^-$.* Define $\mathcal{W}_\rho$ and $\mathcal{W}_\rho^-$ as above in (58), (59).

PROPOSITION 6.20. *For small enough $\rho > 0$ one has*

$$\mathcal{W}_\rho = \overline{\mathcal{V}_\rho \setminus \mathcal{V}_\rho^-}, \qquad \mathcal{W}_\rho^- = \mathcal{W}_\rho \cap \mathcal{V}_\rho^-.$$

Consequently, for small $\rho > 0$ one has $[\mathcal{V}_\rho/\mathcal{V}_\rho^-] = [\mathcal{W}_\rho/\mathcal{W}_\rho^-]$.

Proof. This follows directly from the proof of Proposition 6.18. □

PROPOSITION 6.21. *The pair $(\mathcal{W}_\rho, \mathcal{W}_\rho^-)$ contains $(\mathcal{Z}_\rho, \mathcal{Z}_\rho^-)$ with*

$$\mathcal{Z}_\rho \stackrel{\text{def}}{=} \{u \in \mathcal{W}_\rho \mid w_+(u) = 0\}, \quad \mathcal{Z}_\rho^- \stackrel{\text{def}}{=} \mathcal{Z}_\rho \cap \mathcal{W}_\rho^-$$

as a deformation retract.

Proof. One can write $u \in \mathcal{W}_\rho$ as $u = u^- + u^p + u^+$ and can homotope it to $G_\theta(u) = \mu(\theta)u^- + \nu(\theta)u^p + \theta u^+$, where $\mu(\theta)$, $\nu(\theta) \in \mathbb{R}^+$ are chosen so as to keep $G_\theta(u)$ on $\mathcal{S}_\varepsilon$. □

PROPOSITION 6.22. *$\mathcal{Z}_\rho/\mathcal{Z}_\rho^-$ is homeomorphic with $S^1 \times S^{2p-1}/(S^1 \times \mathrm{pt})$.*

Proof. We can write any $u \in \mathcal{Z}_\rho$ as

$$u(x) = \mathfrak{Re} \textstyle\sum_{k \le p} u_k e^{2\pi i k x/q}, \tag{60}$$

with

$$\textstyle\sum_{k \le p}(1 + k^4)|u_k|^2 = \varepsilon^2. \tag{61}$$

The condition $w_-(u) \le \rho$ is equivalent to

$$\textstyle\sum_{k<p}(1 + k^4)|u_k|^2 \le \rho\varepsilon^2 \tag{62}$$

so that

$$(1 + p^4)|u_p|^2 \ge (1 - \rho)\varepsilon^2.$$

In particular, $u_p \ne 0$ if $\rho < 1$, and we can write

$$u_p = e^{i\theta}\sqrt{\frac{\varepsilon^2 - \sum_{k<p}(1 + k^4)|u_k|^2}{1 + p^4}} \tag{63}$$

with $\theta = \arg u_p$. We have defined a map f from $\mathcal{Z}_\rho$ to $\mathbb{C} \times \mathbb{R} \times \mathbb{C}^{2p-2}$, given by

$$f : u \mapsto (e^{i\theta}, u_0, u_1, \ldots, u_{p-1}).$$

(Recall that u_0 is real, while the other coefficients are complex.)

This map is one-to-one and hence a homeomorphism onto its image. The image is clearly $S^1 \times B^{2p-1}$, where S^1 is the unit circle in $\mathbb{C}$ and B^{2p-1} is the convex ball in $\mathbb{R} \times \mathbb{C}^{p-1}$ given by (62).

The subspace $\mathcal{Z}_\rho^-$ consists of those $u \in \mathcal{Z}_\rho$ for which one has equality in (62), and therefore f maps $\mathcal{Z}_\rho^-$ onto $S^1 \times \partial B^{2p-1}$. We conclude that $\mathcal{Z}_\rho/\mathcal{Z}_\rho^-$ is homeomorphic with $(S^1 \times B^{2p-1})/(S^1 \times \partial B^{2p-1})$ which in turn is homeomorphic with $(S^1 \times S^{2p-1})/(S^1 \times \{\mathrm{pt}\})$. □

6.4. *A long exact sequence relating the $h^I(\mathcal{B})$.* Let $\emptyset \subset J \subset I \subset \{1, \ldots, m\}$ with $J \ne I$ be given, and set $K = I \setminus J$.

Choose a metric $g \in \mathcal{M}_\Gamma(\mathcal{B}; J)$. This metric can be modified to a new metric $\tilde{g} \in \mathcal{M}_\Gamma(\mathcal{B}; I)$ so that g and $\tilde{g}$ coincide on an open neighborhood of the geodesics γ_j, for all $j \in J$.

We can then construct punctured neighborhoods $\mathcal{U}_i \subset \hat{\mathcal{B}}$ of $q_i\gamma_i$ which isolate the $q_i\gamma_i$ for all $i \in I$, and which are so small that curve shortening for g and for $\tilde{g}$ coincide on a neighborhood of $q_i\gamma_i$ in Ω for all $i \in J$.

The indices $h^I(\mathcal{B})$ and $h^J(\mathcal{B})$ are then defined to be the homotopy types of the pointed spaces

$$\mathcal{H}^I(\mathcal{B}) \stackrel{\text{def}}{=} \frac{\hat{\mathcal{B}}}{\mathcal{U}^I \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)}, \quad \mathcal{H}^J(\mathcal{B}) \stackrel{\text{def}}{=} \frac{\hat{\mathcal{B}}}{\mathcal{U}^J \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)},$$

and since $\mathcal{U}^J \subset \mathcal{U}^I$ we have a natural map $\mathcal{H}^J(\mathcal{B}) \to \mathcal{H}^I(\mathcal{B})$ which collapses the set

$$\mathcal{A}^I_J \stackrel{\text{def}}{=} \frac{\mathcal{U}^I \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)}{\mathcal{U}^J \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)} \tag{64}$$

to the base point in $H^I(\mathcal{B})$. Since $\mathcal{U}^I = \mathcal{U}^J \sqcup \mathcal{U}^K$ is a disjoint union, the space $\mathcal{A}^I_J$ in (64) is homeomorphic to

$$\begin{aligned}\frac{\mathcal{U}^I \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)}{\mathcal{U}^J \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)} &= \frac{\mathcal{U}^J \cup \mathcal{U}^K \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)}{\mathcal{U}^J \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)} \\ &= \frac{\mathcal{U}^K}{\mathcal{U}^K \cap \left(\mathcal{U}^J \cup \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g,\varepsilon)\right)} \\ &= \frac{\mathcal{U}^K}{\mathcal{U}^K \cap \mathcal{B}^-} \\ &\cong \bigvee_{k\in K} \left\{\frac{S^1 \times S^{2p_k-1}}{S^1 \times \{\mathrm{pt}\}}\right\}.\end{aligned}$$

Proposition 6.23. *The subset $\mathcal{A}^I_J$ of $\mathcal{H}^J(\mathcal{B})$ is collared.*

Proof. We should have started with neighborhoods $\mathcal{V}_i$, and then chosen $\mathcal{U}_i \subset \mathrm{int}\mathcal{V}_i$. The curve shortening flow then retracts the $\mathcal{V}_i$ into the $\mathcal{U}_i$. $\square$

This proposition implies an isomorphism

$$H_l(\mathcal{H}^J(\mathcal{B}), \mathcal{A}^I_J) \cong H_l(\mathcal{H}^J(\mathcal{B})/\mathcal{A}^I_J) = H_l(\mathcal{H}^I(\mathcal{B}))$$

of relative singular homology groups.

The long exact sequence on homology for the pair $\left(\mathcal{H}^J(\mathcal{B}), \mathcal{A}^I_J\right)$ then gives us the long exact sequence

$$\ldots H_{l+1}(h^I(\mathcal{B})) \xrightarrow{\partial_*} H_l(\mathcal{A}^I_J) \longrightarrow H_l(h^J(\mathcal{B})) \longrightarrow H_l(h^I(\mathcal{B})) \xrightarrow{\partial_*} H_{l-1}(\mathcal{A}^I_J) \ldots \tag{65}$$

from Theorem 1.2.

6.5. *Proof of Theorem* 1.3. We know that not all homology groups of $\mathcal{A}^I_J$ are trivial; so, if $h^I(\mathcal{B})$ is the homotopy type of a point, then the exact sequence implies that $\mathcal{A}^I_J$ and $h^J(\mathcal{B})$ have the same homology groups. Similarly, if $h^J(\mathcal{B})$ happens to be trivial, then $H_l(\mathcal{A}^I_J) \cong H_{l+1}(h^I(\mathcal{B}))$ for all l, so that $h^I(\mathcal{B})$ cannot be trivial.

7. Existence theorems for closed geodesics

7.1. *Proof of Theorem* 1.1. Fix $\Gamma = \{\gamma_1, \dots, \gamma_N\} \subset \Omega$, a relative flat knot type $\mathcal{B} \subset \Omega \setminus \Delta(\Gamma)$, an $I \subset \{1, \dots, m\}$, and a metric $g \in \mathcal{M}_\Gamma(\mathcal{B}; I)$. Assuming that there are no closed geodesics in $\mathcal{B}$ for the metric g we will show that $h^I(\mathcal{B})$ is trivial.

Define $T_* : \hat{\mathcal{B}} \to [0, \infty]$ by

$$T_*(\alpha) = \inf\{t \geq 0 \mid \Phi^t(\alpha) \in \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g, \varepsilon)\}.$$

It was shown in Proposition 6.8 that T_* is continuous. Thus the set

$$W = \{\alpha \in \hat{\mathcal{B}} \mid T_*(\alpha) = \infty\}$$

is closed in $\hat{\mathcal{B}}$.

Choose neighborhoods $\mathcal{U}_i \ni q_i\gamma_i$ with $\mathcal{U}_i = \mathcal{U}_i^\#$, as in §6.2. Let $\mathcal{U} = \cup_{i \in I} \mathcal{U}_i$.

For each $\alpha \in W$ there is a $t_\alpha \in [0, \infty)$ such that $\Phi^{t_\alpha}(\alpha) \in \operatorname{int}\mathcal{U}$, where $\operatorname{int}\mathcal{U}$ is the interior of $\mathcal{U}$ with respect to $\hat{\mathcal{B}}$. Indeed, if $\alpha \in W$ then the entire orbit $\Phi^{[0,\infty)}(\alpha)$ is contained in $\mathcal{B}$. This orbit must converge to some closed geodesic, and by assumption such a geodesic must lie on $\partial\bar{\mathcal{B}}$. That is, the orbit must converge to one of the $q_i\gamma_i$ with $i \in I$.

Continuity of the semiflow implies that some neighborhood $\mathcal{O}_\alpha \ni \alpha$ also gets mapped into $\operatorname{int}\mathcal{U}$ under Φ^{t_α}. Choose a sequence α_n so that the $\mathcal{O}_n = \mathcal{O}_{\alpha_n}$ form a locally finite covering of W. Next let $\mathcal{O} = \cup_n \mathcal{O}_n$ and construct a continuous function $t_0 : \mathcal{O} \to [0, \infty)$ with $t_0(\beta) \geq t_{\alpha_n}$ for all $\beta \in \mathcal{O}_n$. We may assume that $\lim_{\gamma \to \partial\mathcal{O}} t_0(\gamma) = \infty$ (add $\operatorname{dist}(\gamma, \partial\mathcal{O})^{-1}$ to $t_0(\gamma)$ if necessary).

One has $\Phi^{t_0(\beta)}(\beta) \in \operatorname{int}\mathcal{U}$ for all $\beta \in \mathcal{O}$. Moreover,

$$T_\#(\alpha) = \min\bigl(t_0(\alpha), T_*(\alpha)\bigr)$$

defines a continuous everywhere finite function on $\hat{\mathcal{B}}$ which satisfies

$$\Phi^{T_\#(\alpha)}(\alpha) \in \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell(g, \varepsilon) \cup \mathcal{U}.$$

The family of maps

$$F^\theta(\alpha) = \Phi^{\theta T_\#(\alpha)}(\alpha)$$

with $0 \leq \theta \leq 1$ defines a deformation retraction of $(\hat{\mathcal{B}}, \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$ into $(\mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U}, \mathcal{B}^- \cup \hat{\mathcal{B}}^\ell \cup \mathcal{U})$. Thus the index $h^I(\mathcal{B})$ is trivial.

7.2. *Proof of Theorem* 1.4. Let γ be a simple closed geodesic on the sphere S^2. After applying a diffeomorphism we may assume that γ is the equator. We consider p, q satellites of the equator. Thus in the notation we have used so far, we have $\Gamma = \{\gamma_1, \dots, \gamma_N\} = \{\zeta\}$, $N = 1$. The unique curve $\zeta \in \Gamma$ belongs to the boundary of $\mathcal{B}_{p,q}(\zeta)$, and thus $m = 1$. There are two modified Conley indices to be considered, namely $h^\emptyset(\mathcal{B})$ and $h^{\{1\}}(\mathcal{B})$.

To compute the Conley indices $h^I(\mathcal{B}_{p,q}(\zeta))$ for arbitrary $p/q \neq 1$ we use the fact that the indices do not depend on the metric, and consider the standard metric on the usual unit sphere $S^2 \subset \mathbb{R}^3$. For this metric the equator is indeed a closed geodesic, while all geodesics are great circles. In particular, no closed geodesic on the standard sphere is a p, q satellite of the equator. Moreover, the rotation number of the equator is exactly $\rho(\zeta) = 1$.

For $p/q > 1$ we therefore conclude that

$$h^{\{1\}}(\mathcal{B}_{p,q}(\zeta)) = [\text{point}]$$

while for $p/q < 1$ we get

$$h^{\emptyset}(\mathcal{B}_{p,q}(\zeta)) = [\text{point}].$$

By the long exact sequence from Theorem 1.3 we then find that for $p/q > 1$ the index $h^{\emptyset}(\mathcal{B}_{p,q}(\zeta))$ is nontrivial, while for $p/q < 1$ the index $h^{\{1\}}(\mathcal{B}_{p,q}(\zeta))$ is nontrivial.

If one now has another metric g for which the simple closed curve ζ is a geodesic with rotation number $\rho(\zeta, g) > p/q > 1$, then the nontriviality of $h^{\emptyset}(\mathcal{B}_{p,q}(\zeta))$ implies existence of at least one closed geodesic of g which is a (p, q) satellite of ζ. Similarly, if one has $1 > p/q > \rho(\zeta, g)$, then nontriviality of $h^{\{1\}}(\mathcal{B}_{p,q}(\zeta))$ again leads to the same conclusion.

8. Appendices

8.1. *Curve shortening in local coordinates.* Assume g is an $h^{2,\mu}$ metric on M and let $\gamma \in \Omega$ be an $h^{2,\mu}$ curve of length L. Then there exists an $h^{2,\mu}$ local diffeomorphism $\sigma : \mathbb{T} \times (-r, r) \to M$ with $\mathbb{T} = \mathbb{R}/L\mathbb{Z}$ such that $x \mapsto \sigma(x, 0)$ is an arclength parametrization of γ. If γ is a q fold cover, then we may assume that $\sigma(x + L/q, y) \equiv \sigma(x, y)$.

In the local coordinates $\{x, y\}$ the metric g is given by

$$\sigma^* g = E(x, y)(dx)^2 + 2F(x, y)\, dx\, dy + G(x, y)(dy)^2 \tag{66}$$

where E, F, G are $h^{2,\mu}$ functions on $\mathbb{T} \times (-r, r)$.

We now compute the geodesic curvature of the graph of $y = u(x)$ and determine the PDE which is equivalent to curve shortening in the coordinates $\{x, y\}$.

The unit tangent to the graph $\{(x, u(x)) \mid x \in \mathbb{T}\}$ is

$$T = \frac{\partial_x + u_x \partial_y}{|\partial_x + u_x \partial_y|} = \frac{1}{\lambda}(\partial_x + u_x \partial_y)$$

where $\lambda = \sqrt{E + 2F\, u_x + G\, u_x^2}$.

If we write $X \wedge Y$ for $\Omega_g(X, Y)$ where $\Omega_g = (EG - F^2)dx \wedge dy$ is the area form of the metric g, then the geodesic curvature is $\kappa = T \wedge \nabla_T(T)$; i.e.,

$$\kappa = \lambda^{-3}(\partial_x \wedge \partial_y)\left[u_{xx} + P(x,u) + Q(x,u)u_x + R(x,u)u_x^2 + S(x,u)u_x^3\right]$$

where

$$P = \frac{\partial_x \wedge \nabla_{\partial_x}(\partial_x)}{\partial_x \wedge \partial_y}, \qquad Q = \frac{2\partial_x \wedge \nabla_{\partial_y}(\partial_x) + \partial_y \wedge \nabla_{\partial_x}(\partial_x)}{\partial_x \wedge \partial_y},$$
$$R = \frac{\partial_x \wedge \nabla_{\partial_y}(\partial_y) + 2\partial_y \wedge \nabla_{\partial_y}(\partial_x)}{\partial_x \wedge \partial_y}, \qquad S = \frac{\partial_y \wedge \nabla_{\partial_y}(\partial_y)}{\partial_x \wedge \partial_y}.$$

If we now consider a moving family of graphs $y = u(x,t)$, then the normal velocity of this family of curves is given by

$$V = T \wedge (u_t \partial_y) = \lambda^{-1}(\partial_x \wedge \partial_y)u_t$$

so that curve shortening, i.e. $V = \kappa$, is equivalent to

$$u_t = \frac{u_{xx} + P(x,u) + Q(x,u)u_x + R(x,u)u_x^2 + S(x,u)u_x^3}{E(x,u) + 2F(x,u)\,u_x + G(x,u)\,u_x^2}. \tag{67}$$

8.2. *Interpretation of the coefficients P, Q, R, S.* The coefficient $S(x,u)$ is the geodesic curvature of the vertical lines $y =$ constant. If the diffeomorphism σ were obtained by exponentiating normal vectors to the x-axis, as in (13) §2.7, then S would vanish. (However, (13) contains the unit normal vector $\mathbf{N}$ which is only $h^{1,\mu}$, and so the resulting map σ is also only $h^{1,\mu}$ instead of $h^{2,\mu}$.)

The coefficient P is proportional to the geodesic curvature of the curves $u =$ constant. In particular $P(x,0)$ is proportional to the geodesic curvature of the x-axis. The x-axis is a geodesic only if $P(x,0) \equiv 0$.

If the x axis is a geodesic then we can assume after an $h^{2,\mu}$ change of coordinates that $x \mapsto (x,0)$ is a unit speed parametrization of the x-axis, and that on the x-axis the vector ∂_y is a unit normal to the x-axis. In other words, we assume that

$$E(x,0) = 1, F(x,0) = 0, \quad G(x,0) = 1$$

for all x. (Use a Whitney type exension theorem, as in [34, §VI.2.3, Th. 4].)

The derivative $P_y(x,0)$ is then given by

$$\begin{aligned} P_y(x,0) &= \frac{\partial}{\partial y}\left(\frac{\partial_x \wedge \nabla_{\partial_x}(\partial_x)}{\partial_x \wedge \partial_y}\right)_{y=0} \\ &= \partial_x \wedge \nabla_{\partial_y}\nabla_{\partial_x}(\partial_x) \qquad (\text{use } \nabla_{\partial_x}\partial_x = 0, \partial_x \wedge \partial_y = 1 \text{ for } y = 0) \\ &= \partial_x \wedge \left\{\nabla_{\partial_x}\nabla_{\partial_y}(\partial_x) + \mathcal{R}(\partial_y, \partial_x)\partial_x\right\} \qquad (\text{definition of the Riemann tensor}) \\ &= K(x,u) \end{aligned}$$

where K is the Gauss curvature. (This last calculation is the standard derivation of the equation for Jacobi fields.)

On the x-axis we have $\nabla_{\partial_y}(\partial_x) = \nabla_{\partial_x}(\partial_y) = 0$ since ∂_y is a unit normal to a geodesic. We also have $\nabla_{\partial_x}(\partial_x) = 0$ since the x axis is a geodesic. Thus

$$Q(x, 0) = 0.$$

Both R and S are $h^{1,\mu}$ functions of their arguments.

Thus the linearization of (67) at $u = 0$ is

$$u_t = u_{xx} + K(x)u, \tag{68}$$

$K(x) = K \circ \gamma(x)$ being the Gauss curvature along the x-axis.

8.3. *Short time existence and the C^1 local semiflow property.* Equation (67) is of the form

$$u_t = F(x, u, u_x, u_{xx}) \tag{69}$$

where F is a $C^{1,\mu}$ function of its arguments with

$$\lambda^{-1} \leq (1 + p^2)\frac{\partial F(x, u, p, q)}{\partial q} \leq \lambda$$

for some constant λ. It is well-known (perhaps under higher differentiablity assumptions on F) that solutions with initial data $u(\cdot, 0) \in C^{2,\mu}(\mathbb{T})$ exist on a short time interval. We now show that (69) generates a C^1 local semiflow on an open subset of $h^{2,\mu}(\mathbb{T})$.

We may assume in our setting that $F(x, u, p, q)$ is defined for all $(x, u, p, q) \in \mathbb{T} \times \mathbb{R}^3$ with $|u| \leq r$ for some $r > 0$. Let $V \subset C^1(\mathbb{T})$ be defined by

$$V = \{u \in C^1(\mathbb{T}) \mid |u| < r\}.$$

We write $V^{k,\lambda}$ for $V \cap h^{k,\lambda}(\mathbb{T})$.

The PDE (69) is actually quasilinear; i.e., F has the form $F(x, u, p, q) = a(x, u, p)q + b(x, u, p)$ where a and b are $C^{1,\mu}$ in x and u and analytic in $p \in \mathbb{R}$. This implies that the substitution operator $u \mapsto F(x, u, u_x, u_{xx})$ is continuously Fréchet differentiable from $V^{2,\mu}$ to $h^{0,\mu}(\mathbb{T})$. Since the Fréchet derivative of F is the generator of an analytic semigroup in $h^{0,\nu}(\mathbb{T})$ for any $\nu \in (0, \mu)$, we can apply [5, Cor. 2.9] and conclude that (69) generates a C^1 local semiflow on $V^{2,\nu}$ for every $\nu \in (0, \mu)$. This, by definition, means the following:

Continuous local semiflow. The map Φ which maps the initial data u_0 and time t to the solution $u(t)$ at time t is defined on an open subset $\mathcal{D} \subset V^{2,\nu} \times [0, \infty)$ containing $V^{2,\nu} \times \{0\}$ and satisfies

(1) F is continuous,

(2) $F(u_0, 0) = u_0$ for all $u_0 \in V^{2,\nu}$,

(3) If $(u_0, t) \in \mathcal{D}$ then $\{u_0\} \times [0, t] \subset \mathcal{D}$,

(4) If $(u_0, t) \in \mathcal{D}$ and $(F(u_0, t), s) \in \mathcal{D}$ then $(u_0, t+s) \in \mathcal{D}$ and $F(u_0, t+s) = F(F(u_0, t), s)$.

Differentiable local semiflow.[2] For each $t \geq 0$ define $\mathcal{D}_t = \{u \in V^{2,\nu} \mid (u, t) \in \mathcal{D}\}$ and write $\Phi_t(u) = \Phi(u, t)$. Then the map $u \mapsto \Phi_t(u)$ is continuously differentiable from $\mathcal{D}_t$ to $V^{2,\nu}$. Moreover, the Fréchet derivative $d\Phi_t(u)$ is a strongly continuous function of both variables $(u, t) \in \mathcal{D}$, i.e. for any $v_0 \in h^{2,\nu}(\mathbb{T})$ the map $(u, t) \mapsto d\Phi_t(u)v_0$ is continuous from $\mathcal{D}$ to $h^{2,\nu}(\mathbb{T})$. One obtains $d\Phi_t(u)v_0$ by formally linearizing (69); i.e., $v(t) = d\Phi_t(u)v_0$ is the solution of

$$\begin{aligned} &v_t = F_q(x, u, u_x, u_{xx})v_{xx} + F_p(x, u, u_x, u_{xx})v_x + F_u(x, u, u_x, u_{xx})v, \\ &v(\cdot, 0) = v_0(\cdot). \end{aligned} \tag{70}$$

8.4. *Linearization at a closed geodesic.* If the curve γ (the x-axis) is a geodesic so that $u \equiv 0$ is a solution to (67), then $u = 0$ is a fixed point of the local semiflow Φ_t on $V^{2,\nu}$. The semigroup property $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and the chain-rule imply that the linear operators $\{d\Phi_t(0) \mid t \geq 0\}$ form a (C_0) semigroup on $h^{2,\nu}(\mathbb{T})$. Since the linearized equation (70) coincides with (68) the semigroup $\{d\Phi_t(0)\}$ is generated by $\mathcal{A} = \left(\frac{d}{dx}\right)^2 - K \circ \gamma(x)$. ($\mathcal{A}$ is an unbounded operator on h^ν with domain $h^{2,\nu}$ and hence generates a semigroup on h^ν, $h^{2,\nu}$ and any of their interpolation spaces.)

Let the spectrum of $\mathcal{A}$ be

$$\lambda_0 > \lambda_1 \geq \lambda_2 > \cdots > \lambda_{2i-1} \geq \lambda_{2i} > \cdots$$

with corresponding eigenfunctions $\{\varphi_k\}$. For $j \in \mathbb{N}$ we write E_j for span $\{\varphi_0, \ldots, \varphi_{2j}\}$ and E_j^c for the closure in $h^{2,\nu}(\mathbb{T})$ of the span of $\{\varphi_{2j+1}, \varphi_{2j+2}, \ldots\}$. Then E_j and E_j^c are spectral subspaces of the operator $\mathcal{A}$ with $h^{2,\nu}(\mathbb{T}) = E_j \oplus E_j^c$. We let π_j denote the projection of $h^{2,\nu}(\mathbb{T})$ onto E_j along E_j^c.

LEMMA 8.1. *Let $\{u(t) \mid t \geq 0\} \subset V^{2,\nu}$ be an orbit of Φ_t with $\lim_{t\to\infty} u(t) = 0$ in the $h^{2,\nu}(\mathbb{T})$ norm. Then for any $j \in \mathbb{N}$,*

$$\lim_{t\to\infty} \frac{\|\pi_j u(t)\|}{\|u(t)\|} = 0 \text{ or } 1. \tag{71}$$

Here all norms are $h^{2,\nu}(\mathbb{T})$ norms.

The same statement is true for "ancient orbits" $\{u(t) \mid -\infty < t \leq 0\}$ provided all limits are taken for $t \to -\infty$.

---

[2] We repeat these definitions here because there seems to be no consensus on what a differentiable local semiflow should be. In particular Amann [1], [2] does not include or prove strong continuity of $d\Phi_t$ at $t = 0$.

Proof. Once one gets away from the PDE and considers $\{u(n) \mid n = 1, 2, \cdots\}$ as an orbit of the time-one map Φ_1 the proof is completely standard.

The map $\Phi_1 : \mathcal{D}_1 \to V^{2,\nu}$ is C^1, and its Fréchet derivative is given by $d\Phi_1(0) = e^{\mathcal{A}}$, a compact operator with spectrum e^{λ_i}, $i \in \mathbb{N}_0$. One can find equivalent norms $|||\cdot|||$ on E_j and E_j^c so that

$$|||e^{\mathcal{A}}v||| \geq e^{\lambda_{2j}}|||v|||, \qquad \forall v \in E_j,$$
$$|||e^{\mathcal{A}}v||| \leq e^{\lambda_{2j+1}}|||v|||, \qquad \forall v \in E_j^c.$$

Suppose now that

$$\limsup_{n\to\infty} \frac{|||\pi_j u(n)|||}{|||u(n)|||} > 0.$$

Then for some $\varepsilon > 0$ and any $r > 0$ there exists a large n_* such that $\frac{|||\pi_j u(n_*)|||}{|||u(n_*)|||} > \varepsilon$ and $|||u(n)||| < r$ for all $n \geq n_*$.

We can now write $\Phi_1(u) = \mathcal{M}(u)u$ where

$$\mathcal{M}(u) = \int_0^1 d\Phi(\theta u)d\theta.$$

Since Φ_1 is C^1 we have $|||\mathcal{M}(u) - e^{A}||| < \sigma(|||u|||)$ where $\sigma(r) \searrow 0$ as $r \searrow 0$.

If one splits $u = v \oplus v^c \in E_j \oplus E_j^c$, as well as $\bar{u} = M(u)u = \bar{v} \oplus \bar{v}^c$, and if one assumes $|||u||| \leq r$, then

$$\begin{aligned}
|||\bar{v}||| &= |||\pi_j \mathcal{M}(u)(v \oplus v^c)||| \\
&\geq |||\pi_j e^{\mathcal{A}}(v \oplus v^c)||| - \sigma(r)|||u||| \\
&\geq e^{\lambda_{2j}}|||v||| - \sigma(r)|||u||| \\
&\geq \left(e^{\lambda_{2j}} - \sigma(r)\right)|||v||| - \sigma(r)|||v^c|||.
\end{aligned}$$

Similarly one finds

$$|||\bar{v}^c||| \leq \left(e^{\lambda_{2j+1}} + \sigma(r)\right)|||v^c||| + \sigma(r)|||v|||.$$

If one also assumes that $|||v||| \geq \varepsilon|||v^c|||$ then

$$\frac{|||\bar{v}|||}{|||\bar{v}^c|||} \geq \frac{\left(e^{\lambda_{2j}} - \sigma(r)\right)\varepsilon - \sigma(r)}{e^{\lambda_{2j+1}} + \sigma(r) + \varepsilon\sigma(r)}. \tag{72}$$

Since $e^{\lambda_{2j}} > e^{\lambda_{2j+1}}$ one can choose $1 < \vartheta < e^{\lambda_{2j}-\lambda_{2j+1}}$. For sufficently small $r > 0$ one concludes from (72) that

$$\frac{|||\bar{v}|||}{|||\bar{v}^c|||} \geq \vartheta\varepsilon.$$

Inductive application of this estimate shows that for $u(k) = v(k) \oplus v^c(k)$ one has

$$\frac{|||v(n_* + i)|||}{|||v^c(n_* + i)|||} \geq \varepsilon(n_* + i) = \vartheta^i\varepsilon$$

as long as $1 + \varepsilon(n_* + i) < \frac{\delta}{2\sigma(r)}$.

Thus if $\limsup_{n\to\infty} \|\pi_j u(n)\|/\|u(n)\| > 0$, then

$$\liminf_{n\to\infty} \|\pi_j v(n)\|/\|v^c(n)\| \geq \delta/2\sigma(r),$$

with $r > 0$ arbitrarily small. Hence $\|\pi_j u(n)\|/\|u(n)\| \to 1$ as $n \to \infty$.

Having established the alternative (71) along a sequence $n \nearrow \infty$ we now assume that

$$\limsup_{t\nearrow\infty} \|\pi_j u(t)\|/\|u(t)\| > \varepsilon > 0.$$

Then for any $r > 0$ there is a $t_0 > 0$ such that $\|u(t)\| \leq r$ for $t \geq t_0$ and $\|\pi_j u(t_0)\|/\|u(t_0)\| > \varepsilon$.

The previous arguments imply that $\lim_{n\nearrow\infty} \|\pi_j u(t_n)\|/\|u(t_n)\| = 1$, where $t_n = t_0 + n$.

Splitting $u(t) = v(t) \oplus v^c(t)$ as before we have $\|v(t_n)\| = o\big(\|v^c(t_n)\|\big)$ for $n \nearrow \infty$. To estimate $v(t)$ and $v^c(t)$ for $t \in (t_n, t_{n+1})$ we write

$$u(t_n + \theta) = d\Phi_\theta(u(t_n)) = \mathcal{M}_\theta(u(t_n))u(t_n),$$

where $\mathcal{M}_\theta(u) = \int_0^1 d\Phi_\theta(su)ds$.

Since Φ is a differentiable semiflow the map $(\theta, u) \mapsto \mathcal{M}_\theta(u)$ is strongly continuous. Hence, for small enough $r > 0$ the operators $\{\mathcal{M}_\theta(u) \mid 0 \leq \theta \leq 1, \|u\| \leq r\}$ are uniformly bounded. Since E_{2j} is finite dimensional, the map $(\theta, u) \mapsto \mathcal{M}_\theta(u)|_{E_{2j}}$ is norm continuous. In particular, there is a $\tau(r) > 0$ with $\tau(r) \searrow 0$ for $r \searrow 0$, such that

$$\|\mathcal{M}_\theta(u)|_{E_{2j}} - e^{\theta\mathcal{A}}|_{E_{2j}}\|_{\mathrm{L}(E_{2j}, h^{2,\nu})} \leq \tau(r)$$

if $\|u\| \leq r$ and $\theta \in [0, 1]$.

We have the following estimates:

$$\begin{aligned}
\|v(t_n + \theta)\| &\geq \|\pi_j \mathcal{M}_\theta(u(t_n))v(t_n)\| - \|\pi_j \mathcal{M}_\theta(u(t_n))v^c(t_n)\| \\
&\geq \|\pi_j e^{\theta\mathcal{A}} v(t_n)\| - \|\pi_j(\mathcal{M}_\theta(u(t_n)) - e^{\theta\mathcal{A}})v(t_n)\| \\
&\quad - \|\pi_j \mathcal{M}_\theta(u(t_n))v^c(t_n)\| \\
&\geq e^{\theta\lambda_{2j}}\|v(t_n)\| - \tau(r)\|v(t_n)\| - o(1)\|v(t_n)\|
\end{aligned}$$

(use $\|v^c(t_n)\| = o(\|v(t_n)\|)$). Also

$$\begin{aligned}
\|v^c(t_n + \theta)\| &\leq \|\pi_j^c \mathcal{M}_\theta(u(t_n))v(t_n)\| - \|\pi_j^c \mathcal{M}_\theta(u(t_n))v^c(t_n)\| \\
&\leq \tau(r)\|v(t_n)\| + o(1)\|v(t_n)\| \\
&= o(1) \cdot \|v(t_n)\|
\end{aligned}$$

which together imply $\|v^c(t_n + \theta)\| = o(\|v(t_n + \theta)\|)$ as $n \nearrow \infty$, uniformly in $\theta \in [0, 1]$. $\square$

LEMMA 8.2 (Notation as in Lemma 8.1). *For any solution* $\{u(t) \mid t \geq 0\}$ *of curve shortening which converges to* $u = 0$ *there exists a* $j \in \mathbb{N}$ *such that*

$$\lim_{t\to\infty} \frac{\|\pi_j u(t)\|}{\|u(t)\|} = 1.$$

In particular one has $\|u(t)\| \geq Ce^{-\zeta t}$ *for some* $\zeta < \infty$.

Proof. If the limit were 0 for all j then the solution $u(t)$ would approach $u = 0$ faster than any exponential, and so we must prove the lower bound $\|u(t)\| \geq Ce^{-\zeta t}$.

There is a standard approach for proving exponential *lower* bounds on decay in heat equations due to Agmon (see [19, §2.18, p.181]) which is used to prove backward uniqueness results. This approach would work here, but it would require us to differentiate the functions $a(x, u, p)$ in the PDE twice, thereby forcing us to consider metrics g on M with at least three derivatives. In order not to use more than just $g \in h^{2,\mu}$ we follow the less standard approach from the appendix in [3] which applies to semilinear equations.

To rewrite curve shortening as a semilinear equation we study the evolution of the curvature as a function of renormalized arclength. Let $\gamma : \mathbb{R}/\mathbb{Z} \times [0, t_*) \to M$ be a normal parametrization (i.e. $\partial_t\gamma \perp \partial_x\gamma$) of a solution of curve shortening. Write $L(t)$ for length at time t, let P_t be the point $\gamma(0, t)$ (so that P_t moves with velocity perpendicular to the curve always) and define the normalized arclength coordinate ς of any point $Q = \gamma(x, t)$ on γ_t by

$$\varsigma(x,t) = \frac{1}{L(t)} \int_0^x |\partial_x\gamma(\xi, t)| d\xi = \frac{1}{L(t)} \int_{P_t}^Q ds. \tag{73}$$

We also introduce a new time variable related to t via

$$\tau = \int_0^t \frac{dt}{L(t)^2}.$$

PROPOSITION 8.3. *The curvature* κ, *as a function of* τ *and* ς, *satisfies*

$$\kappa_\tau = \kappa_{\varsigma\varsigma} + L(\tau)^2 \left\{ J[\kappa]\kappa_\varsigma + (K \circ \gamma)\kappa + \kappa^3 \right\} \tag{74}$$

where

$$J[\kappa] = \int_0^\varsigma \kappa^2 d\varsigma - \varsigma \int_0^1 \kappa^2 d\varsigma.$$

Proof. A straightforward calculation begins with differentiating (73) with respect to t to get

$$\frac{\partial\varsigma}{\partial t} = \varsigma \int_0^1 \kappa^2 d\varsigma - \int_0^\varsigma \kappa^2 d\sigma.$$

Then the chain rule

$$\left(\frac{\partial\kappa}{\partial t}\right)_{x=\text{const}} = \left(\frac{\partial\kappa}{\partial t}\right)_{\varsigma=\text{const}} + \frac{\partial\kappa}{\partial\varsigma}\frac{\partial\varsigma}{\partial t}$$

after some simplification leads to (74). □

Since the limiting geodesic $u = 0$ has positive length the new and old time variables t and τ are roughly proportional, so it suffices to establish an exponential lower bound for the solution in the τ variable.

The equation (74) is semilinear, and can be written as

$$\kappa_\tau = \mathcal{A}\kappa + \mathcal{R}(\tau)\kappa$$

where $\mathcal{A} = (\partial_\varsigma)^2 - L_0^2 K_0(x)$, with $K_0(x) = K(\sigma(x, 0))$, is the Gauss curvature on the x axis, and L_0 is the length of the x-axis. The “remainder” operator $\mathcal{R}(\tau)$ is

$$\mathcal{R}(\tau) = L(\tau)^2 J[\kappa(\tau)] \frac{\partial}{\partial\varsigma} + \kappa^2 + \bigl(K \circ \gamma_\tau - K \circ \gamma_\infty\bigr).$$

This operator is bounded from the Sobolev space $W^{1,2}(\mathbb{R}/\mathbb{Z})$ to $L^2(\mathbb{R}/\mathbb{Z})$. If we assume that $\|\pi_j u(t)\|/\|u(t)\| \to 0$ for all j then the coefficients in $\mathcal{R}$ decay faster than any exponential $e^{-\zeta\tau}$ and thus the operator norm of $\mathcal{R}(\tau)$ from $W^{1,2}(\mathbb{R}/\mathbb{Z})$ to $L^2(\mathbb{R}/\mathbb{Z})$ also tends to zero.

The eigenvalues of the self-adjoint operator $\mathcal{A}$ on L^2 grow like n^2, so the n^{th} gap in the spectrum of $\mathcal{A}$ has length proportional to n. This is exactly enough for the argument in [3, Appendix] and we can conclude that no solution of curve shortening can approach a geodesic at a faster than exponential rate. $\square$

University of Wisconsin-Madison, Madison, Wisconsin
E-mail address: angenent@math.wisc.edu

References

[1] H. Amann, Quasilinear evolution equations and parabolic systems, *Trans. Amer. Math. Soc.* **293** (1986), 191–227.

[2] ———, Dynamic theory of quasilinear parabolic equations. I. Abstract evolution equations, *Nonlinear Anal.* **12** (1988), 895–919.

[3] S. B. Angenent, The Morse-Smale property for a semi-linear parabolic equation, *J. Differential Equations* **62** (1986), 427–442.

[4] ———, The periodic orbits of an area preserving twist map, *Comm. Math. Phys.* **115** (1988), 353–374.

[5] ———, Nonlinear analytic semiflows, *Proc. Royal Society of Edinburgh* **115A** (1990), 91–107.

[6] ———, Parabolic equations for curves on surfaces. I. Curves with p-integrable curvature, *Ann. of Math.* **132** (1990), 451–483.

[7] ———, Parabolic equations for curves on surfaces. II. Intersections, blow-up, and generalized solutions, *Ann. of Math.* **133** (1991), 171–215.

[8] ———, On the formation of singularities in the curve shortening flow, *J. Differential Geometry* **33** (1991), 601–633.

[9] S. B. Angenent, Recent results in mean curvature flow, in *Recent Advances in Partial Differential Equations* (El Escorial, 1992), 1–18, *RAM Res. Appl. Math.* **30**, Masson, Paris, 1994.

[10] ———, Inflection points, extatic points and curve shortening, in *Hamiltonian Systems Systems with Three or More Degrees of Freedom* (S'Agaró, 1995), 3–10, *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* **533**, Kluwer Acad. Publ., Dordrecht, 1999.

[11] ———, The zero set of a solution of a parabolic equation, *J. Reine Angew. Math.* **390** (1988), 79–96.

[12] S. B. Angenent and B. Fiedler, The dynamics of rotating waves in reaction diffusion equations, *Trans. Amer. Math. Soc.* **307** (1988), 545–568.

[13] V. I. Arnol'd, *Topological Invariants of Plane Curves and Caustics, University Lecture Series* **5**, A.M.S., Providence, RI, 1994.

[14] ———, *Mathematical Methods of Classical Mechanics*, *Grad. Texts in Math.* **60**, Springer-Verlag, New York, 1978.

[15] G. D. Birkhoff, *Dynamical Systems*, *A.M.S. Colloq. Publ.* **IX**, A.M.S., Providence, RI, 1966.

[16] ———, Surface transformations and their dynamical applications, *Acta Math.* **43** (1922), 1–199.

[17] C. Conley, *Isolated Invariant Sets and The Morse Index*, *CBMS Regional Conf. Series in Math.* **38** (1978), A.M.S. Providence, RI.

[18] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[19] A. Friedman, *Partial Differential Equations*, Holt, Rinehart and Winston, Inc., New York, 1969.

[20] M. Gage, An isoperimetric inequality with applications to curve shortening, *Duke Math. J.* **50** (1983), 1225–1229.

[21] ———, Curve shortening makes convex curves circular, *Invent. Math.* **76** (1984), 357–364.

[22] M. Gage and R. S. Hamilton, The heat equation shrinking convex plane curves, *J. Differential Geom.* **23** (1986), 69–96.

[23] M. Grayson, *The heat equation shrinks embedded plane curves to round points,* J. *Differential Geom.* **26** (1987), 285–314.

[24] ———, Shortening embedded curves, *Ann. of Math.* **129** (1989), 71–111.

[25] R. S. Hamilton, Isoperimetric estimates for the curve shrinking flow in the plane, in *Modern Methods in Complex Analysis* (Princeton, NJ, 1992), 201–222, *Ann. of Math. Studies* **137**, Princeton Univ. Press, Princeton, NJ, 1995.

[26] G. Huisken, A distance comparison principle for evolving curves, *Asian J. Math.* **2** (1998), 127–133

[27] R. Johnson and J. Moser, The rotation number for almost periodic potentials, *Comm. Math. Phys.* **84** (1982), 403–438.

[28] W. Klingenberg, Lectures on closed geodesics, *Grundlehren der mathematischen Wissenschaften*, **230** (1978), Springer-Verlag, New York.

[29] H. Matano, Nonincrease of the lap-number of a solution for a one-dimensional semilinear parabolic equation, *J. Fac. Sci. Univ. Tokyo Math.* **29** (1982), 401–441.

[30] J. N. Mather, Amount of rotation about a point and the Morse index, *Comm. Math. Phys.* **94** (1984), 141–153.

[31] K. Nickel, Gestaltaussagen über Lösungen parabolischer Differentialgleichungen, *J. Reine Angew. Math.* **211** (1962), 78–94.

[32] J. A. Oaks, Singularities and self-intersections of curves evolving on surfaces, *Indiana Univ. Math. J*. **43** (1994), 959–981.

[33] H. Poincaré, Sur les lignes geodesiques des surfaces convexes, *Trans. Amer. Math. Soc.* **6** (1905), 237–274.

[34] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, *Princeton Math. Series* **30**, Princeton Univ. Press, Princeton, NJ, 1970.

[35] C. Sturm, Mémoire sur une classe d'équations à différences partielles, *J. de Mathématiques Pures et Appliquées* **1** (1836), 373–444.

(Received December 4, 2000)