

THE CLOSED GEODESIC PROBLEM FOR COMPACT RIEMANNIAN 2-ORBIFOLDS

JOSEPH E. BORZELLINO AND BENJAMIN G. LORICA

In this paper it is shown that any compact Riemannian 2-orbifold whose underlying space is a (compact) manifold without boundary has at least one closed geodesic.

Introduction.

In this paper, we examine the question of the existence of a smooth closed geodesic on Riemannian 2-orbifolds. Roughly speaking a Riemannian orbifold is a metric space locally modelled on quotients of Riemannian manifolds by finite groups of isometries. It turns out that Riemannian orbifolds inherit a natural stratified length space structure and are sufficiently well-behaved locally so that one may apply both techniques of Alexandrov geometry and geometric analysis to extend standard results about Riemannian manifolds to Riemannian orbifolds. The 2-orbifolds we consider in this paper are orbifolds whose underlying space is a manifold *without* boundary. One can think of such Riemannian orbifolds as 2-manifolds with some distinguished singular cone points, whose neighborhoods are isometric to a quotient of the 2-disc with some metric by a cyclic group of finite order fixing the center of the disc. The 2-orbifolds we consider fall into two categories which we will handle with different techniques. The first case is when the underlying space of the orbifold is simply connected (in the usual topological sense), that is, the underlying space of the orbifold is the 2-sphere S^2 . This class of orbifolds contains the set of all orientable *bad* 2-orbifolds, namely those that do not arise as a quotient of S^2 with some metric by a finite group of isometries acting properly discontinuously. These bad 2-orbifolds are examples of what are commonly referred to as teardrops and footballs. The second class of 2-orbifolds are those whose underlying space is not simply connected in the usual sense. The basic reference for orbifolds is [T], while a more Riemannian viewpoint is taken in [B1]. Many of the results on Riemannian orbifolds that we will use have appeared in published form in [B2].

Before we state and discuss our results for Riemannian orbifolds, we would like to recall the methods and ideas used to prove the classical theorem of Fet and Lyusternik [FL]: On any compact Riemannian manifold there

exists at least one closed geodesic. The essential tool in proving this result, in an elementary way, is to develop a process of curve-shortening. This process is commonly attributed to Birkhoff [Bi]. The idea here is, given a continuous map of say the unit interval into our manifold M , to divide the interval into small subintervals so that the endpoints of the curve restricted to any subinterval have the property that there exists a unique minimal geodesic connecting the two endpoints. That such a subdivision exists follows from compactness of M , since then one finds a uniform lower bound on the injectivity radius at any point of M . By replacing the given curve by the minimal geodesic connecting such endpoints one constructs a new “broken” geodesic homotopic to the original of length less than or equal to that of the original. Now one iterates this process by joining those endpoints that correspond to the midpoints of the previous subintervals with the minimal geodesic connecting them. In this way one generates at each stage a new broken geodesic of shorter length, which is homotopic to the original. It is worth mentioning that if one is interested in applying this process to closed curves, namely maps of $S^1 = [0, 1]/\{0, 1\}$ into M that at each stage this process yields a closed broken geodesic *freely* homotopic to the original. Now by compactness, essentially the Arzela–Ascoli theorem, one can find a subsequence of these broken geodesics which converge, and in fact will converge to a geodesic. We refer to [K, Section 3.7] for the details.

An alternate approach to the Fet–Lyusternik theorem is to apply techniques from the calculus of variations on Hilbert manifolds, see [S, Chapter 8]. If M is a closed Riemannian manifold, the space of $H^1(S^1, M)$ curves is a manifold modelled on a Hilbert space. The geodesics correspond precisely to the critical points of an appropriate energy functional defined on the space $H^1(S^1, M)$. The energy functional satisfies the famous Palais–Smale compactness condition, the main analytic tool needed in proving the existence of critical points. While this approach has natural aesthetic advantages over the polygonal approximation approach mentioned above, Bott [Bo] notes that the use of global analysis does not appear to be essential for any aspect of the geodesic problem on closed manifolds. The use of infinite dimensional manifolds, however, is of fundamental importance in the study of other geometric variational problems such as the study of minimal surfaces and Yang–Mills theory.

Part of the proof of the main result of this paper requires that we work on compact manifolds with boundary. It is not clear to the authors how one should choose to construct a suitable structure on $H^1(S^1, M)$ in the case that ∂M is non-empty. It is for this reason that we adopt an approach similar in spirit to the polygonal approximation construction outlined above.

In trying to generalize the result of Fet and Lyusternik to Riemannian

footballs one must overcome the following difficulty: there is not a uniform lower bound on the injectivity radius at points in a compact Riemannian orbifold. This follows from a result of the first author [B2, Proposition 15] where it is shown, for example, that a minimal geodesic cannot enter and leave the singular set. As a result the injectivity radius of a non-singular point is bounded above by its distance to the singular set, and hence no uniform bound is possible (unless of course the singular set is empty and M is a Riemannian manifold). We now state our main result.

Theorem 1. *Let O be a compact Riemannian 2-orbifold whose underlying space is a (compact) manifold without boundary. Then O has at least one closed geodesic.*

Remark 2. Riemannian orbifolds carry the structure of a length space (or inner metric space). By *geodesic* we mean a path in the orbifold which is locally length minimizing. This agrees with the definition of geodesic for general length spaces. When working with orbifolds, however, we should point out that it is common to define a geodesic as a path that lifts locally to a geodesic. These two notions are related but are not equivalent.

We would like to thank J. Hass for useful conversations regarding this work. We would also like to thank P. Petersen for reading an earlier version of this paper and suggesting improvements of the original results.

Review of the Curve-Shortening Process.

In this section, we let X denote a smooth compact Riemannian manifold with (or without) boundary. Then there exists a real number i_0 such that any two points $p, q \in X$ with $d(p, q) < i_0$ can be joined by a unique minimal geodesic which depends continuously on the two points.

We define a curve-shortening process along the lines of that described in [GZ]. Let $\gamma : S^1 = [0, 1]/\{0, 1\} \rightarrow X$ be a closed curve in X . Assume that γ is parametrized proportional to arclength. Denote by L the length of γ . Let m be an integer such that $L/m < i_0$. Divide the curve γ into m equal segments each of length L/m , by the division points $q_0, q_1, \dots, q_{m-1}, q_m$. Now replace each arc $\widehat{q_i q_{i+1}}$ by the unique minimal geodesic $\overline{q_i q_{i+1}}$ joining q_i to q_{i+1} of length $< i_0$. This replaces γ by the m -sided closed geodesic polygon

$$\gamma' = \overline{q_0 q_1} \cup \overline{q_1 q_2} \cup \dots \cup \overline{q_{m-1} q_m}.$$

Note that the length of γ' is strictly smaller than the length of γ unless $\gamma' = \gamma$. Now take the m midpoints of the segments of γ' . Successive midpoints are at distance $< i_0$ from each other and hence can be joined by a unique minimal

geodesic. This produces a new m -sided geodesic polygon γ'' . The process described above is to be one iteration of the curve-shortening process. We denote γ'' as $\Phi(\gamma)$. Continuing inductively, we see that at each stage we have produced a new curve homotopic to and of length not longer than the curve of the previous stage.

The Non-Simply Connected Case.

We consider in this section the case when the underlying space of the orbifold is not simply connected (in the usual sense). The argument presented here is a modified version of an argument which originally appeared in the first author's Ph.D. thesis [B1].

As usual, we denote the singular set by Σ . Let \mathcal{C} be a non-trivial free homotopy class. Let $\ell = \inf \{L(c) \mid c \in \mathcal{C}\}$. Then $\ell > 0$, for if there exists a sequence $\{c_n\} : [0, 1] \rightarrow O$ such that $L(c_n) \rightarrow 0$ with c_n parametrized proportional to arc length, then by the Arzela-Ascoli theorem some subsequence of $\{c_n\}$ converges to a continuous curve c . Since length is lower-semicontinuous, we have $L(c) = 0$ which implies c is a constant path. But O is locally simply connected, hence $c_n \sim c$ for large n which is a contradiction. Thus, $\ell > 0$. Now choose a sequence $\{c_n\}$ such that $L(c_n) < \ell + \frac{1}{n}$. Then as before, $\{c_n\}$ form an equicontinuous family with $\{c_n(t)\}$ bounded. Hence $c_n \rightarrow c$ a continuous curve in \mathcal{C} . We have $L(c) \leq \ell$ and hence by definition of ℓ , $L(c) = \ell$. We now show that c is a closed geodesic. If $c \cap \Sigma = \emptyset$, then c is a closed geodesic, for otherwise it could be shortened locally. If $c \cap \Sigma \neq \emptyset$, then c cannot be minimal in any neighborhood of the singular set which follows from [B2, Proposition 15]. Hence we can get a shorter curve $\tilde{c} \sim c$ with $\tilde{c} \cap \Sigma = \emptyset$, which contradicts construction of c . This completes the proof in the non-simply connected case.

The Simply Connected Case.

We are considering the situation when the underlying space of O is the 2-sphere S^2 . We will split our argument for this situation into two cases. The first case will be where O has no more than two singular points (teardrops and footballs) and the other when O has at least 3 singular points.

The Teardrop and Football case. Let O compact Riemannian 2-orbifold with two or fewer singular points. Denote by p and q , the singular points of O . If O has only one singular point p , choose q to be any point of say maximal distance from p . If O has no singular points, that is, O is a smooth 2-sphere, choose p and q realizing the diameter of O . We will refer to p and q as the singular points of O (whether or not they are truly singular). Denote by Σ the singular set $\{p\} \cup \{q\}$. For $0 < \delta < d(p, q)/3$ denote by O_δ the set of

points $x \in O$ such that $d(x, \Sigma) \geq \delta$. Then O_δ is the manifold with boundary $S^1 \times I$. By [ABB, Theorem 5], every point in O_δ possesses a neighborhood which is convex in the sense that any two points in the neighborhood may be joined by a unique geodesic entirely contained in the neighborhood. In fact, it is not hard to see that such a neighborhood may be chosen to be a metric ball. Hence, by compactness there exists a positive real number $r_{\text{convex}} > 0$, the convexity radius, for which any metric ball of radius at most r_{convex} is convex.

Fix $\delta_0 = d(p, q)/3$. Choose δ small enough so that the boundary circles x_δ and y_δ of O_δ have length $< r_{\text{convex}}$, the convexity radius of O_{δ_0} , and so that $\Phi_\delta^k(x_\delta) \subset B_p(\frac{1}{3}d(p, q))$ and $\Phi_\delta^k(y_\delta) \subset B_q(\frac{1}{3}d(p, q))$, where Φ_δ^k denotes the k -th iterate of the curve shortening process in O_δ . We also require that δ is so small that the lengths $L(x_\delta)$ and $L(y_\delta)$ are non-increasing as $\delta \rightarrow 0$. Note that this can be done since neighborhoods of the singular points are asymptotically Euclidean cones (or smooth Euclidean discs).

By applying curve shortening to x_δ , y_δ , we produce two closed geodesics (in O_δ) $x_\delta \rightarrow x_\infty \subset B_p(\frac{1}{3}d(p, q))$ and $y_\delta \rightarrow y_\infty \subset B_q(\frac{1}{3}d(p, q))$. Now foliate $O - \{p, q\}$ by circles such that for $\delta' \leq \delta$ these foliating circles are exactly the distance spheres from p and q . Using this foliation we can produce a path $F : [0, 1] \rightarrow \Lambda O_\delta$, where ΛO_δ denotes the loop space of O_δ , with $F(0) = x_\infty$ and $F(1) = y_\infty$. To see this, we construct the path F as follows: Start at x_∞ . Run the curve shortening process backwards to x_δ . Now follow the foliating circles until you reach y_δ . Now apply curve shortening to go from y_δ to y_∞ . Also we have that

$$\sup_t E(F(t)) \leq M < \infty$$

where E denotes the energy functional on ΛO_δ . It is easy to see that by the construction of the path F that the constant M can be chosen to be independent of δ .

Recall (see [Bo], for example) that there is a finite dimensional approximation to the subset $E^{-1}[0, 2M) \subset \Lambda O_\delta$. That is to say, there is a finite dimensional manifold O_δ^r homotopy equivalent to $E^{-1}[0, 2M)$. In fact, $O_\delta^r \subset O_\delta \times O_\delta \times \cdots \times O_\delta$, and O_δ^r contains all closed geodesics of energy $< 2M$. In particular, we have that x_∞ and y_∞ are contained in O_δ^r .

Since $F : [0, 1] \rightarrow \Lambda O_\delta$, $F(0) = x_\infty$, $F(1) = y_\infty$, and $\sup_t E(F(t)) \leq M$, the set

$$\Omega = \left\{ \Gamma : [0, 1] \rightarrow O_\delta^r \mid \Gamma(0) = x_\infty, \Gamma(1) = y_\infty, \sup_t E(\Gamma(t)) \leq M \right\}$$

is non-empty. Define

$$c_\delta = \inf_{\Gamma \in \Omega} \sup_{0 \leq t \leq 1} E(\Gamma(t)).$$

It follows that c_δ is a critical value for $E : O_\delta^* \rightarrow \mathbb{R}$. To see this, suppose to the contrary that there exists $\varepsilon > 0$ such that the set $E^{-1}[c_\delta - \varepsilon, c_\delta + \varepsilon]$ contains no critical points. Choose $\Gamma_\varepsilon \in \Omega$ such that

$$\sup_{0 \leq t \leq 1} E(\Gamma_\varepsilon(t)) < c_\delta + \varepsilon.$$

Now apply the curve-shortening flow Φ_s to the path $\Gamma_\varepsilon(t)$. By curve-shortening flow we mean the continuous flow that can be constructed from the discrete curve shortening process. This can be found in [Bo]. Observe that for all $t \in [0, 1]$ there exists $s(t)$ such that $E(\Phi_{s(t)}(\Gamma_\varepsilon(t))) < c_\delta - \varepsilon$, and also that Φ_s fixes both x_∞ and y_∞ since they are already closed geodesics. By compactness, we have that

$$\sup_{0 \leq t \leq 1} s(t) = \bar{s} < \infty$$

and thus

$$E(\Phi_{\bar{s}}(\Gamma_\varepsilon(t))) \leq c_\delta - \varepsilon.$$

Moreover we have that $\Phi_{\bar{s}} \circ \Gamma_\varepsilon \in \Omega$, but this contradicts the definition of c_δ , so c_δ is a critical value. Let γ_δ be the critical point associated with the critical value c_δ . We have that $E(\gamma_\delta) \leq M$. It also follows by the triangle inequality and the homotopic essentiality of γ_δ that the length $L(\gamma_\delta) \geq \min\{\text{convexity radius of } O_{\delta_0}, \frac{1}{3}d(p, q)\}$. In particular, the length of γ_δ is bounded below by a constant independent of δ .

If as $\delta \rightarrow 0$, γ_δ lies entirely within the interior $\text{int}(O_\delta)$ for any fixed stage δ then we are done as then γ_δ is a closed geodesic which lies entirely outside the singular set. If this does not happen then we may assume that for all δ , $q_\delta \in \gamma_\delta \cap \partial O_\delta$. Without loss of generality we may assume that $q_\delta \rightarrow q \in \Sigma$, and that $\gamma_\delta \rightarrow \gamma_0$ and that $q \in \gamma_0$. By results in [AA] and [ABB], we know that γ_δ is differentiable at q_δ . Thus, the tangent vectors $\dot{\gamma}_\delta$ and $-\dot{\gamma}_\delta$ make an angle of π at q_δ . Suppose that the angle of the curve γ_0 at q is $< \pi$. Then it must follow that the tangential (to ∂O_δ) components of the second derivatives

$$\|\tan(\nabla_{\dot{\gamma}_\delta} \dot{\gamma}_\delta)\| \rightarrow \infty$$

as $\delta \rightarrow 0$, but this contradicts the length minimizing property of γ_δ at q_δ . See [ABB]. Hence the angle of the curve γ_0 at q must be π , and this a contradiction unless q is non-singular.

If q is non-singular, we show that in fact γ_0 is a geodesic. The only problem is that γ_0 might not be minimizing across q . If this is the case, choose points u, v on γ_0 which straddle q and are close to q . Let $L(u, v)$ denote the length of the segment along γ_0 joining u to v . Then since γ_0 is

not minimal, there exists a minimal geodesic σ joining u to v . Assume that σ lies in O_{δ_σ} and choose $\varepsilon > 0$ such that $L(u, v) = L(\sigma) + \varepsilon = d(u, v) + \varepsilon$. Now choose $\delta < \delta_\sigma$ such that $d(u_\delta, u) < \varepsilon/4$, $d(v_\delta, v) < \varepsilon/4$ and such that $|L(u, v) - d(u_\delta, v_\delta)| < \varepsilon/4$, where u_δ, v_δ are points on γ_δ . Then we have

$$\begin{aligned} d(u_\delta, v_\delta) &\leq d(u_\delta, u) + d(u, v) + d(v, v_\delta) < \frac{\varepsilon}{2} + L(u, v) - \varepsilon \\ &< \frac{\varepsilon}{2} + d(u_\delta, v_\delta) + \frac{\varepsilon}{4} - \varepsilon < d(u_\delta, v_\delta) \end{aligned}$$

which is a contradiction, and hence γ_0 must have been minimal through q .

Thus, in either case, we have produced the desired closed geodesic of positive length. The proof is now complete for the simply connected case with two or fewer singular points.

The Case of More Than Two Singular Points. In this situation we are assuming that our orbifold O is the 2-sphere with more than two singular points. We will use the notation of the previous section. Let the singular set be $\Sigma = \{p_1, \dots, p_n\}$ with $n \geq 3$. Let $r_0 = \frac{1}{2} \min\{d(p_i, p_j), i \neq j\}$. For $\delta < r_0$ we have that $O_\delta = D^2 - \bigcup_{1 \leq i \leq n-1} B_{p_i}(\delta)$, where D^2 is the 2-disc, and $B_{p_i}(\delta)$ denotes the metric δ -ball centered at p_i . In particular, the fundamental group $\pi_1(O_\delta)$ is the free group on $n-1$ generators $\{a_1, \dots, a_{n-1}\}$. Let $\gamma \in [a_1 a_2^{-1}] \in \pi_1(O_\delta)$. Let γ_δ be the limit of the curve-shortening process applied to γ in O_δ . Since γ is not homotopically trivial, the length of γ_δ is > 0 . We claim that the length $L(\gamma_\delta)$ of γ_δ is bounded below as $\delta \rightarrow 0$. If this is not the case, then it follows that for δ small enough that $\gamma_\delta \subset B_{p_i}(\frac{1}{3}r_0)$, for otherwise γ_δ would be entirely contained in some convex ball of O_δ , and hence homotopically trivial. Also, if $\gamma_\delta \subset B_{p_i}(\frac{1}{3}r_0)$, then it must follow that γ is freely homotopic to a_i^m for some m , which is a contradiction. Thus, the length of γ_δ is bounded below as $\delta \rightarrow 0$. By arguing as in (the end of) the previous section, we can conclude that for some $\delta > 0$, γ_δ must have been a closed geodesic missing the boundary of O_δ , and hence is a closed geodesic in O . This completes the proof of the simply connected case, and hence finishes the proof of Theorem 1.

References

- [AA] R. Alexander and S. Alexander, *Geodesics in Riemannian Manifolds With Boundary*, Indiana U. Math J., **30** (1981), 481-488.
- [ABB] S. Alexander, I. Berg and R. Bishop, *The Riemannian Obstacle Problem*, Illinois J. Math., **31** (1987), 167-184.
- [B1] J. Borzellino, *Riemannian Geometry of Orbifolds*, Ph.D thesis, University of California, Los Angeles 1992.
- [B2] ———, *Orbifolds of Maximal Diameter*, Indiana U. Math. J., **42** (1993), 37-53.

- [Bi] G.D. Birkhoff, *Dynamical Systems with Two Degrees of Freedom*, Trans. Amer. Math. Soc., **18** (1917), 199-300.
- [Bo] R. Bott, *Lectures on Morse Theory, Old and New*, Bull. of the AMS, **7**(2) (1982), 331-358.
- [FL] A. Fet and L. Lyusternik, *Variational Problems on Closed Manifolds*, Dokl. Akad. Nauk. SSSR, **81** (1951), 17-18 [Russian].
- [GZ] H. Gluck and W. Ziller, *Existence of Periodic Motions of Conservative Systems*, in Seminar on Minimal Submanifolds, Princeton Univ. Press, Princeton 1983.
- [K] W. Klingenberg, *Riemannian Geometry*, Walter deGruyter, New York 1982.
- [S] J.T. Schwartz, *Nonlinear Functional Analysis*, with an additional chapter by H. Karcher, Gordon and Breach, 1969.
- [T] W. Thurston, *The Geometry and Topology of 3-Manifolds*, Lecture Notes, Princeton University Math. Dept., 1978.

Received April 13, 1994.

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CA 90095-1555
E-mail address: borzelli@math.ucla.edu

AND

INSTITUTE FOR MATHEMATICAL SCIENCES AND APPLICATIONS
CALIFORNIA STATE UNIVERSITY, MONTEREY BAY
SEASIDE, CA 93955-8001
E-mail address: ben_lorica@qmbbridge.calstate.edu