CHARACTERIZATIONS OF THE ROUND TWO-DIMENSIONAL SPHERE IN TERMS OF CLOSED GEODESICS

LEE KENNARD AND JORDAN RAINONE

Abstract. The question of whether a closed Riemannian manifold has infinitely many geometrically distinct closed geodesics has a long history. Though unsolved in general, it is well understood in the case of surfaces. For surfaces of revolution diffeomorphic to the sphere, a refinement of this problem was introduced by Borzellino, Jordan-Squire, Petrics, and Sullivan. In this article, we quantify their result by counting distinct geodesics of bounded length. In addition, we reframe these results to obtain a couple of characterizations of the round two-sphere.

All closed Riemannian manifolds contain a closed geodesic. If the manifold is not simply connected, any length-minimizing representative of a nontrivial homotopy class is a closed geodesic. In the simply connected case, this is already a nontrivial result.

A more difficult question is whether there exist infinitely many closed geodesics. To avoid over-counting, one considers two geodesics *geometrically distinct* if their images are distinct. This brings us to the well known question of whether there exist infinitely many geometrically distinct closed geodesics. In this article, we restrict our attention to surfaces, but we refer the reader to Oancea [12, Chapter 2] for a survey and a guide to the literature on the problem.

For surfaces with genus $g \geq 1$, one uses the infinitude of the fundamental group and a length minimization argument to construct infinitely many geometrically distinct closed geodesics. For the torus, it follows that the number of such geodesics of length at most ℓ grows quadratically in ℓ (see Berger [2, Chapter XII.5.A]). For $g \geq 2$, Katok proved that this number actually grows exponentially in ℓ (see Remark 0.3 below).

In the remaining case, when the surface is the sphere, this question was only answered affirmatively in the 1990s by Bangert and Franks [1, 5] (cf. [2] and Hingston [6]). Hingston then proved a quantified version of this result (see [7]): Given any metric on $\mathbb{S}^2$, the number of geometrically distinct closed geodesics of length at most ℓ is asymptotically at least $c\ell/\log \ell$ for some constant $c > 0$.

In this article, we consider refinements of these results. As motivation, consider a surface of revolution. Each profile curve connecting the poles extends to a closed geodesic. In particular, the results of Bangert–Franks and Hingston are trivial in this setting. On the other hand, all of these geodesics are in some sense the same. This motivates the following definition: For a closed Riemannian manifold M, we say that two geodesics on M are *strongly geometrically distinct* if there is no isometry taking the image of one to the image of the other.

For metrics with finite isometry group, one has immediate analogues of the results above. For metrics with infinite symmetry, it is unclear whether there exist infinitely many strongly geometrically distinct geodesics. For example, the constant curvature metric on $\mathbb{S}^2$ has only one closed geodesic in this sense. In [3], Borzellino et al. prove that all surfaces of revolution diffeomorphic to $\mathbb{S}^2$, except for the round spheres, have infinitely many strongly geometrically distinct geodesics. Our main result is a quantification of this result, as well as a straightforward observation that it extends to all closed, orientable surfaces with continuous (equivalently infinite) symmetry.

Main Theorem. *Let M be an orientable, compact surface with infinite isometry group. Let $N(\ell)$ denote the number of strongly geometrically distinct closed geodesics on M of length less than or equal to ℓ. One of the following occurs:*

(1) *M is isometric to a round sphere, and $N(\ell) = 1$ for all sufficiently large $\ell > 0$.*
(2) *There is a constant $c > 0$ such that $N(\ell) \geq c\ell^2$ for all sufficiently large $\ell > 0$.*

We make a few remarks.

Remark 0.1. In the non-orientable case, one applies the theorem to the orientable double cover to obtain an analogous characterization of the real projective plane with constant curvature.

Remark 0.2. It is well known that a closed, orientable surface M can have infinite isometry group only if M is diffeomorphic to $\mathbb{S}^2$ or the torus T^2 (see Lemma 1.1). In the latter case, a simple extension of a standard argument shows the Main Theorem holds. However the argument we provide for $\mathbb{S}^2$ carries over with little effort to the case of T^2, so we include it in Section 3 for completeness.

Remark 0.3. For a compact surface M with genus $g \geq 2$, the isometry group is finite, so $N(\ell)$ is related to the number $n(\ell)$ of geometrically distinct closed geodesics on M of length at most ℓ by the following relation:

$$N(\ell) \leq n(\ell) \leq CN(\ell),$$

where C denotes the number of elements in the isometry group. Hence asymptotics on $n(\ell)$ imply asymptotics on $N(\ell)$, up to multiplicative constant. For a metric on M with constant curvature -1, Margulis showed that the function $n(\ell)$ is asymptotic to ce^ℓ/ℓ for some constant c, i.e., the quotient $n(\ell)/(ce^\ell/\ell) \to 1$ as $\ell \to \infty$ (see Margulis [11], cf. Katok [9, Section 1]). In particular, $n(\ell) \leq e^\ell$ for all sufficiently large ℓ. On the other hand, Katok showed that, for any metric on M with the same area as the constant curvature -1 metric,

$$\liminf_{\ell\to\infty} \log(n(\ell))/\ell \geq 1,$$

with equality if and only if the metric has constant curvature -1 (see Katok [8], cf. Berger [2, Chapter XII.5.B]). As a consequence, for the case of non-constant curvature, there exists a constant $a > 1$ such that $n(\ell) \geq e^{a\ell}$ for all sufficiently large ℓ. Hence for both $\mathbb{S}^2$ and surfaces of genus $g \geq 2$, there is a sense in which the constant curvature metric is characterized by having the fewest closed geodesics. We do not know whether the constant curvature metrics on T^2 have a similar characterization.

Consider now a metric on $\mathbb{S}^2$ with infinite isometry group. The metric takes the form $ds^2 + h(s)^2 d\theta^2$ and one can check that the arguments in Borzellino et al. for a surface of revolution carry over to this slightly more general case to show that infinitely many strongly geometrically distinct closed geodesics exist, i.e., $\lim_{\ell\to\infty} N(\ell) = \infty$. In Section 2, we summarize their argument and supplement it where needed to prove the claimed lower bound on the growth rate of $N(\ell)$.

Before starting the proof, we point out that this theorem, combined with the work of Hingston and Katok, immediately implies the following:

Corollary. *Let M be an orientable, compact surface. Either M is isometric to a round sphere and $N(\ell) = 1$ for all sufficiently large $\ell > 0$, or there exists a constant $c > 0$ such that $N(\ell) \geq c\ell/\log \ell$ for all sufficiently large $\ell > 0$.*

Acknowledgements. This project began as part of the Summer Undergraduate Research Fellowship program in the College of Creative Studies at UCSB. The second author is grateful for the support provided by this program. The first author was partially supported by NSF grants DMS-1045292 and DMS-1404670. Both authors would like to thank Wolfgang Ziller for helpful comments in the preparation of this article.

1. Preliminaries on Lie group actions

In this section, we gather some results on isometric actions by Lie groups that are required for the proofs. We summarize the results here:

Lemma 1.1. *If M is a closed, orientable Riemannian manifold of dimension two with infinite isometry group G, then the identity component $G_0 \subseteq G$ contains a circle S^1, and one of the following occurs:*

1. *M is isometric to a round $\mathbb{S}^2$ and $\dim(G) = 3$.*
2. *M is diffeomorphic to $\mathbb{S}^2$ but not isometric to a round $\mathbb{S}^2$, $\dim(G) = 1$, and the fixed-point set of S^1 is a pair of isolated points.*
3. *M is diffeomorphic to a torus, and the fixed-point set of S^1 is empty.*

In particular, M cannot have genus $g \geq 2$.

To prove this lemma, suppose M is a closed Riemannian manifold of dimension two with infinite isometry group G. A theorem of Myers and Steenrod states that G is a compact Lie group (see Kobayashi [10, Chapter II, Section 1]). Let $G_0 \subseteq G$ denote the identity component. By compactness, G has only finitely many components. Since G is infinite, this implies that G_0 has positive dimension. In particular, the maximal torus theorem implies that G_0 contains a circle S^1.

This circle acts isometrically on M, and its fixed-point set

$$F = \{p \in M \mid e^{it}(p) = p \text{ for all } e^{it} \in \mathrm{S}^1\}$$

equals the zero set of the associated Killing field X on M defined by $X(p) = \frac{d}{dt}\big|_{t=0} (e^{it}(p))$. Moreover, F consists of isolated points, and the number of these points equals the Euler characteristic of M (see [10, Chapter II, Theorems 5.3 and 5.5]). Since the Euler characteristic of M equals $2-2g$ where g is the genus, it follows either that

M is diffeomorphic to $\mathbb{S}^2$ and F is a pair of isolated points or that M is diffeomorphic to T^2 and F is empty.

It suffices to show that $\dim(G) = 3$ if and only if M is a round $\mathbb{S}^2$, and that $\dim(G) = 2$ only if M is diffeomorphic to T^2. Regarding the first of these claims, we note that a round $\mathbb{S}^2$ has isometry group $\mathrm{O}(3)$, which is three-dimensional. Conversely, it is a classical fact that, if the isometry group of a compact two-manifold is three-dimensional, then M is either $\mathbb{S}^2$ or the real projective plane $\mathbb{RP}^2$ equipped with a metric of constant curvature (see [10, Chapter II, Theorem 3.1]). If, moreover, M is orientable, as in Lemma 1.1, then we conclude that M is isometric to a round $\mathbb{S}^2$.

Suppose now that $\dim(G) = 2$. The only compact, connected, two-dimensional Lie group is the two-torus, so $G_0 = T^2$ (see Bröcker–tom Dieck [4, page 169]). Since G_0 acts effectively on M and has the same dimension as M, it follows that G_0 acts transitively on M and hence that the Gauss curvature is constant. By the Gauss–Bonnet theorem and the fact that the genus $g \leq 1$, either M is a round $\mathbb{S}^2$ or a flat T^2. In the first of these cases, we have $\dim(G) = 3$, a contradiction to the assumption that $\dim(G) = 2$. Hence M is isometric to a torus with constant zero curvature.

2. Proof of Main Theorem for the sphere

Assume that M is a Riemannian manifold diffeomorphic to $\mathbb{S}^2$ with infinite isometry group. Let $\{p, q\} \subseteq M$ denote the fixed point set of this circle action according to Lemma 1.1. Choose a minimal geodesic c from p to q. By rescaling the metric if necessary, assume that c is defined on $[0, \pi]$ and that $c(0) = p$ and $c(\pi) = q$. There exists a smooth function $h : (0, \pi) \to (0, \infty)$ and an isometric covering map

$$\begin{array}{rccc} \sigma\colon & \big((0,\pi)\times\mathbb{R}, ds^2 + h(s)^2 d\theta^2\big) & \longrightarrow & M \setminus \{p,q\} \\ & (s,\theta) & \mapsto & e^{i\theta}\cdot c(s), \end{array}$$

where the dot denotes the action of the circle element $e^{i\theta}$ on $c(s)$. Since M is smooth at $p = c(0)$ and $q = c(\pi)$, we conclude that the extended function $h : [0, \pi] \to \mathbb{R}$ satisfies $h(0) = h(\pi) = 0$ and $h'(0) = -h'(\pi) = 1$ (see [13, Chapter 1, Section 3.4]). The strategy now is to follow the proof in Borzellino et al. [3], which covers the case of a surface of revolution. Note that, for a surface of revolution, $h(s)$ represents one coordinate of a unit-speed curve in the plane and hence satisfies the condition that $|h'(s)| \leq 1$ (see Petersen [13, Chapter 1, Section 3.3, Example 18]). Although we are considering a more general class of surfaces, the arguments of Borzellino et al. extend to our situation. We summarize the proof here since our strategy is simply to supplement it, as needed, in order to prove the Main Theorem.

In the coordinates induced by σ, the geodesic equations are

$$\begin{aligned} s''(t) &= h(s(t))h'(s(t))\theta'(t)^2, \\ \theta''(t) &= -2\frac{h'(s(t))}{h(s(t))}s'(t)\theta'(t). \end{aligned}$$

The meridians, $\gamma(t) = \sigma(t, \theta_0)$, satisfy these equations and extend to closed geodesics passing through both poles, p and q. Since θ_0 is arbitrary, we have by uniqueness that meridians are the only geodesics that pass through the poles. In the rest of this section, we consider those geodesics that do not pass through the poles. Since σ defines an isometric covering map onto $M \setminus \{p, q\}$, we can write a geodesic $\gamma(t)$ as $\sigma(s(t), \theta(t))$ for smooth functions $s : \mathbb{R} \to (0, \pi)$ and $\theta \colon \mathbb{R} \to \mathbb{R}$. For example, the parallels given by $\gamma(t) = \sigma(s_0, t/h(s_0))$ are closed geodesics provided that $h'(s_0) = 0$. Another example of a geodesic is provided in Figure 1.

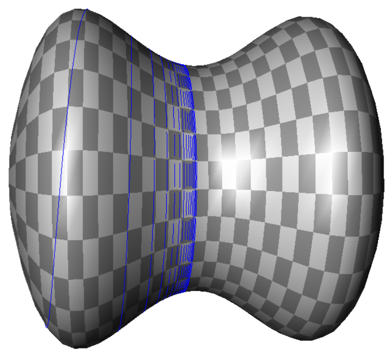

FIGURE 1. A geodesic asymptotic to a parallel. The surface is $\mathbb{S}^2$ equipped with a rotationally symmetric metric.

An important consequence of the geodesic equations is Clairaut's relation. This states that, for each non-meridian geodesic γ, there exists a constant $c_\gamma > 0$ such that

$$h(s(t)) \cos \alpha(t) = c_\gamma,$$

where $\alpha(t)$ is the angle between $\gamma'(t)$ and the coordinate vector field σ_θ at $\gamma(t)$. Since the cosine function is bounded, $h(s(t))$ cannot go to zero, hence any non-meridian curve has its s–coordinate bounded by some interval

$$[s_0(\gamma), s_1(\gamma)] = [\inf s(t), \sup s(t)] \subseteq (0, \pi).$$

Further analysis shows the following.

Lemma 2.1 (Clairaut)**.** *For $a \in (0, \pi)$, let γ_a be a unit-speed geodesic starting with s–coordinate a and initial direction $\gamma'(0)$ in the θ-direction. One of the following occurs:*

1. **parallel:** *$h'(a) = 0$, and $s(t) = a$ for all t.*
2. **asymptotic:** *$h'(a) > 0$ (resp. < 0) and there exists $b = b(a) > a$ (resp. $< a$) such that $h'(b) = 0$ and $s(t) \to b$ as $t \to \infty$.*
3. **oscillating:** *$h'(a) > 0$ (resp. < 0) and there exists $b = b(a) > a$ (resp. $< a$) such that $h'(b) < 0$ (resp. > 0) and $s(t)$ oscillates between a and b, achieving these extremal values at integral multiples of some time, denoted $T(a)$.*

According to this result, we refer to the parameter $a \in (0, \pi)$ as parallel, asymptotic, or oscillating. Following [3, Proposition 3.1], we let $U \subseteq (0, \pi)$ denote the subset consisting of oscillating $a \in (0, \pi)$ for which $h'(a) > 0$ and $h'(b(a)) < 0$, where $b(a) = \inf\{b > a \mid h(b) = h(a)\}$. Geometrically, the s–coordinate of γ_a oscillates between a and $b(a)$. It follows that $U \subseteq (0, \pi)$ is an open set and that the function $a \mapsto b(a)$ on U is smooth. Indeed, this function is given by h composed with a local inverse of h, and so it is smooth by the inverse function theorem. Figure 2 indicates the region U for a function $h(s)$ corresponding to the dumbbell shape from Figure 1.

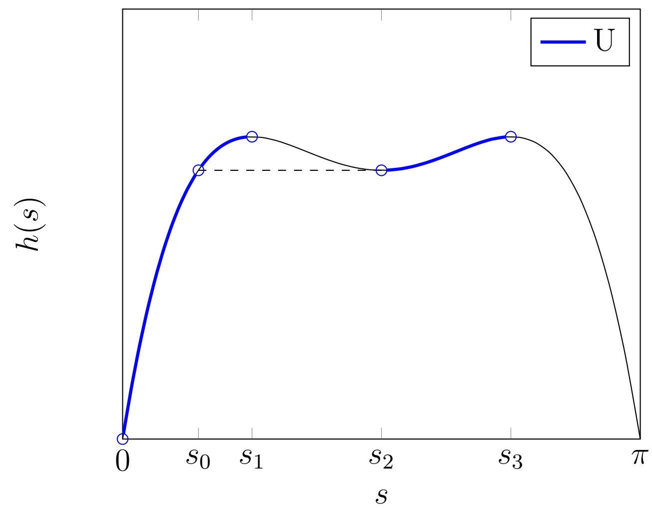

FIGURE 2. Example of function $h(s)$ corresponding to a surface of revolution with the shape of a dumbbell, as in Figure 1. Here, s is the arclength coordinate. The value $a = s_0$ corresponds to an asymptotic geodesic as in Lemma 2.1, and the values $a \in \{s_1, s_2, s_3\}$ correspond to parallel geodesics. The blue region is U, the set of oscillating values of a for which $h'(a) > 0$.

For each $a \in U$, let $\gamma_a(t) = \sigma(s(t), \theta(t))$ be as in Lemma 2.1 and define

$$\begin{aligned} R(a) &= 2\int_0^{T(a)} \theta'(t)dt, \\ L(a) &= 2T(a) = 2\int_0^{T(a)} 1dt, \end{aligned}$$

where $T(a)$ is the time referred to in the third conclusion of Lemma 2.1. This defines two functions $R : U \to \mathbb{R}$ and $L : U \to \mathbb{R}$. The geometric interpretation of these functions is as follows. The quantity $2T(a)$ denotes the time required for a geodesic starting at $s = a$ and parallel to σ_θ to have its s–coordinate go to $b(a)$ and back to a. We call this a "full trip". It then follows by symmetry that $R(a)$ and $L(a)$ denotes the total rotation and length of the geodesic on a full trip. In [3], the authors prove that $R(a)$ is a continuous function of a. For our purposes, we also need that $L(a)$ is continuous.

Lemma 2.2. *The functions $L, R : U \to \mathbb{R}$ are continuous.*

Proof. The proofs for R and L are similar, so we only prove it for L. Fix $a \in U$. Choose a non-trivial interval $[a_1, a_2] \subseteq U$ containing a on which $h' \geq c_1 > 0$. We prove now that L is continuous on $[a_1, a_2]$.

To do this, we rewrite expression for $L(a)$. First, the unit-speed condition implies that $1 = |\gamma_a'(t)|^2 = s'(t)^2 + h(s(t))^2\theta'(t)^2$. Since $s(t)$ is increasing from $t = 0$ to $t = T(a)$, this implies

$$s'(t) = \sqrt{1 - h(s(t))^2\theta'(t)^2}.$$

Next, the second geodesic equation implies that $\frac{d}{dt}\left(h(s(t))^2\theta'(t)\right) = 0$. As a result, $h(s(t))^2\theta'(t)$ equals a constant C. At $t = 0$, the unit-speed condition implies that $\theta'(0) = 1/h(s(0)) = 1/h(a)$, so we have that $C = h(a)$. Putting this together, we obtain

$$s'(t) = \sqrt{1 - h(a)^2/h(s(t))^2}.$$

Finally, we use this expression in order apply the change of variables $s = s(t)$ to the integral $L = 2\int_0^{T(a)} dt$. This gives us the expression

$$L = 2\int_a^{b(a)} \frac{ds}{\sqrt{1 - h(a)^2/h(s)^2}}.$$

Regarding the right-hand side as a function of a, we may write $L(a) = 2\int_a^{b(a)} l(a,s)ds$, where $l(a,s)$ is given by $h(s)/\sqrt{h(s)^2 - h(a)^2}$. This integral is improper at both endpoints, so we proceed by proving the following two claims:

(1) For all sufficiently small $\delta > 0$, $L_\delta(a) = 2\int_{a+\delta}^{b(a)-\delta} l(a,s)ds$ is smooth.
(2) The functions L_δ converge uniformly to L on $[a_1, a_2]$.

The first claim follows from the Leibniz integral rule since $l(a,s)$ is a smooth function on the set $\{(a,s) | a \in [a_1, a_2], a + \delta \leq s \leq b(a) - \delta\}$. To prove the second claim, it suffices to prove that $\int_a^{a+\delta} l(a,s)ds \to 0$ and $\int_{b(a)-\delta}^{b(a)} l(a,s)ds \to 0$ uniformly in $a \in [a_1, a_2]$ as δ goes to 0. These claims are proven similarly, so we only prove the first. The second only requires the additional fact that $b(a)$ depends smoothly on a.

Observe that $l(a,s)$ is non-negative and bounded above as

$$l(a,s) = \frac{h(s)}{\sqrt{h(s)^2 - h(a)^2}} \leq \frac{1}{2c_1}\frac{2h(s)h'(s)}{\sqrt{h(s)^2 - h(a)^2}}.$$

Integrating this expression and applying the change of variables $y = h(s)^2 - h(a)^2$, we conclude that

$$\int_a^{a+\delta} l(a,s)ds \leq \frac{1}{2c_1}\int_0^{h(a+\delta)^2 - h(a)^2} \frac{dy}{\sqrt{y}} = \frac{\sqrt{h(a+\delta)^2 - h(a)^2}}{c_1}.$$

Since h is smooth and hence uniformly continuous on $[0, \pi]$, this last quantity converges to 0 uniformly in a as $\delta \to 0$. This completes the proof. $\square$

We proceed to the proof of the Main Theorem, that the number $N(\ell)$ of strongly geometrically distinct closed geodesics grows quadratically in ℓ. The idea is to show, for all large $\ell > 0$, that a large number of values of a exist such that $a \in U$, $R(a) = 2\pi\frac{p}{q}$

for some rational $\frac{p}{q}$, and $L(a) \leq \ell/q$. These three conditions imply that any choice of γ_a as in Lemma 2.1 is oscillating, closes up after q "full trips", and is a closed geodesic with length at most ℓ.

First, we dispose of the case where the isometry group G satisfies $\dim(G) \neq 1$. By Lemma 1.1, we have $\dim(G) = 3$ and that M is a round sphere. In this case, the isometry group is O(3) or SO(3), and every unit-speed geodesic can be carried to any other by an isometry, so $N(\ell) = 1$ for all ℓ larger than $2\pi r$, where $1/r^2$ is the Gauss curvature of M. This completes the proof of the Main Theorem in this case.

We assume from now on that $\dim(G) = 1$. As a result, the identity component $G_0 \subseteq G$ equals the circle group. By compactness, G has only finitely many components. In particular, for each oscillating value of a as above, at most finitely many other such values result in geodesics that are not strongly geometrically distinct from γ_a. This issue results in a multiplicative factor (equal to the number of components in the isometry group) in our estimates. Since the Main Theorem involves an unknown multiplicative constant anyway, we simply assume, without loss of generality, that the isometry group equals the circle.

The proof is carried out in three cases, which are based roughly on the setup in [3]. One key step is to prove that there exists an asymptotic geodesic if h has more than one critical point. This actually need not be the case. Indeed, a capped cylinder provides a counterexample, since every critical point is a local maximum and hence not a limiting value of an asymptotic geodesic. This problem is easy to fix, however, by breaking the proof into cases as follows.

Lemma 2.3. *If h has infinitely many critical points, then $N(\ell) = \infty$ for all sufficiently large $\ell > 0$.*

Proof. If $h'(a) = 0$, then $\gamma_a(t) = \sigma(a, t/h(a))$ is a closed geodesic of length $2\pi h(a)$. Moreover, the image of γ_a maps to itself under any isometry, so distinct values of a yield strongly geometrically distinct closed geodesics. The result follows since h is bounded on $[0, \pi]$. □

Lemma 2.4. *If h has finitely many critical points, and R is locally constant, then $N(\ell) = \infty$ for all sufficiently large $\ell > 0$.*

Proof. In this case, the argument in [3, Corollaries 4.4 and 4.5] is valid since the critical points are isolated. Indeed, first suppose that h has more than one critical point (as in Figure 2). The arguments there show that M has an asymptotic geodesic and hence that R is unbounded on U. However Lemma 2.1 and the assumptions of this lemma imply that R takes on only finitely many values, so this is a contradiction. Assume instead that h has a unique critical point, s_0 (as in Figure 3 below). It follows as in [3, Corollary 5.4] that $U = (0, s_0)$ and that $R(a) = \lim_{a' \to 0} R(a') = 2\pi$ for all $a \in (0, s_0)$. But L is continuous on $(0, s_0)$ and hence on $[\frac{s_0}{3}, \frac{s_0}{2}]$, so there exist infinitely many strongly geometrically distinct closed geodesics of length at most L_0, where $L_0 = \max\{L(s) \mid s \in [\frac{s_0}{3}, \frac{s_0}{2}]\} < \infty$. □

Lemma 2.5. *If h has finitely many critical points and R is not locally constant, then there exists a constant $c > 0$ such that $N(\ell) \geq c\ell^2$ for all sufficiently large $\ell > 0$.*

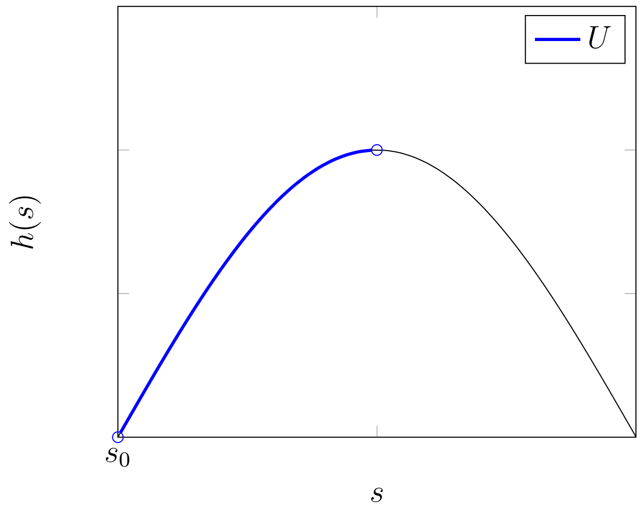

FIGURE 3. An example of a profile curve $h(s)$ with a unique critical point. As in Figure 2, s is the arclength parameter and U is the set of oscillating s–values a for which $h'(a) > 0$.

Proof. Choose a closed interval $I' \subseteq U$ that is mapped by R to some non-trivial interval $I \subseteq \mathbb{R}$. Let $2\pi\frac{p}{q} \in I$. Each $a \in U$ that is mapped by R to $2\pi\frac{p}{q}$ corresponds to a closed geodesic of length $qL(a)$. Since L is continuous on I', this length is at most qL_0, where L_0 is the maximum value of L on I'. This length is at most ℓ if and only if $q \leq \lfloor \ell/L_0 \rfloor$. To estimate $N(\ell)$ from below, it suffices to count the number of rationals $\frac{p}{q} \in \frac{1}{2\pi}I$ with $q \leq \lfloor \ell/L_0 \rfloor$. By Lemma 2.6 below, there is a constant c' such that the number of such rationals is at least $c'\left(\lfloor \ell/L_0 \rfloor\right)^2$ for all sufficiently large ℓ. Taking $c = \frac{1}{2}c'/L_0^2$, we conclude that $N(\ell) \geq c\ell^2$ for all sufficiently large $\ell > 0$. $\square$

As indicated in the previous proof, it suffices to prove the following counting lemma.

Lemma 2.6. *Inside any connected, non-trivial interval $I \subseteq \mathbb{R}$, there exist constants $c > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there are at least cn^2 rational numbers in I with denominator at most n.*

Proof. The proof uses Farey fractions. Let F_n denote the set of rationals a/b written in reduced form such that $0 \leq a \leq b \leq n$. It is easy to see that the number of elements in F_n satisfies

$$|F_n| = 1 + \sum_{k=1}^{n} \phi(k),$$

where $\phi(k)$ is the Euler totient function, given by the number of integers $1 \leq i \leq k$ coprime to k. According to Walfisz [14],

$$\sum_{k=1}^{n} \phi(k) = \frac{3}{\pi^2}n^2 + \mathrm{O}\left(n\left(\log n\right)^{2/3}\left(\log\log n\right)^{4/3}\right).$$

In particular, it follows that constants $c_1 > 0$ and $n_0 > 0$ exist such that $|F_n| > c_1 n^2$ for all $n \geq n_0$.

The idea now is to inject F_n into I in a controlled way. First, it is clear that the conclusion of the lemma holds for I if and only if it holds for $\{1+i \mid i \in I\}$. Hence, we assume without loss of generality that $I \not\subseteq (-\infty, 0]$. Choose positive integers a and b such that I contains the interval $\left[\frac{a}{b}, \frac{a+1}{b}\right]$. Set $c = \frac{1}{2}\left(\frac{c_1}{b^2}\right)$, and choose $n_0 \geq n_1$ such that $\lfloor n/b \rfloor \geq n_1$ and $c_1\left(\frac{n}{b} - 1\right)^2 > cn^2$ for all $n \geq n_0$. We claim that $n \geq n_0$ implies that the number of rationals $x \in I$ with denominator at most n is at least cn^2.

To do this, consider the injection $F_{\lfloor n/b \rfloor} \to I$ given by $x \mapsto \frac{a+x}{b}$. Note that the rationals in the image of this map have denominator at most n. Hence the total number of rationals in I with denominator at most n is at least the order of $F_{\lfloor n/b \rfloor}$. For all $n \geq n_0$, this order is at least $c_1\left(\lfloor n/b \rfloor\right)^2$, which in turn is greater than cn^2. $\square$

This completes the proof of the Main Theorem in the case where M is a sphere.

3. Proof of Main Theorem for the torus

Assume now that M is diffeomorphic to the torus and has infinite isometry group. In this case, there exists an isometric covering map from

$$\sigma : (\mathbb{R} \times \mathbb{R}, ds^2 + h(s)^2 d\theta^2) \to M,$$

where $h : \mathbb{R} \to \mathbb{R}$ is some smooth, positive, and periodic function on $\mathbb{R}$, as in Figure 4. To fix notation, we perform a global scaling so that the period is π.

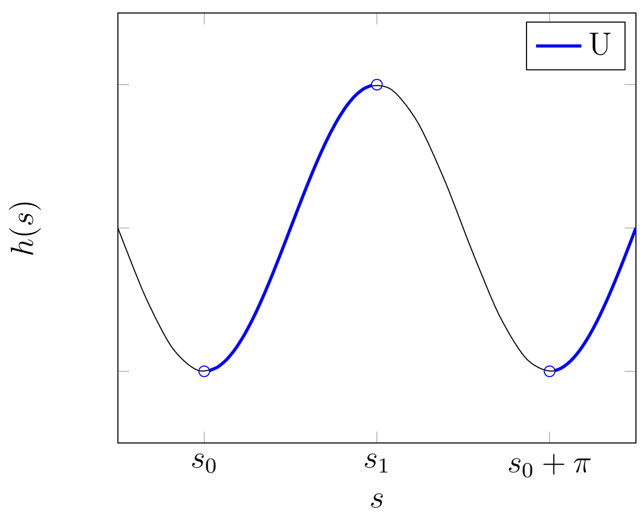

Figure 4. Example of function $h(s)$ corresponding to a torus of revolution. Here, s is the arclength coordinate. The s–values congruent to s_0 or s_1 modulo π correspond to parallel geodesics. The blue region labeled U is, by analogy with the sphere case, the set of oscillating s–values a such that $h'(a) > 0$.

As with the case where M is diffeomorphic to $\mathbb{S}^2$, we obtain the same geodesic equations and Clairaut relation. However, Lemma 2.1 does not hold since it is possible for geodesics to have the property $|s(t)| \to \infty$ as $t \to \infty$. Indeed, this is the case for meridians. As a substitute, we make the following easy observation.

Lemma 3.1. *The π–periodic function $h : \mathbb{R} \to \mathbb{R}$ has at least one of the two following properties:*

(1) *(non-isolated case) There exist infinitely many critical points in $(0, \pi)$.*
(2) *(asymptotic case) There exists an isolated local minimum at some $s_0 \in \mathbb{R}$.*

In the first case of the lemma, it follows that $N(\ell) = \infty$ for all $\ell \geq 2\pi \max(h)$. In the second case, it follows as in the case where M is a sphere that the rotation function $R(a)$ is unbounded. One can imagine why this happens if $h(s)$ is as in Figure 4, since $R(a) \to \infty$ as $a \to s_0$ from the right. Given that $R(a)$ is unbounded, it follows that $R(a)$ is not locally constant and hence that $N(\ell) \geq c\ell^2$ asymptotically in ℓ for some constant $c > 0$. This concludes the proof in this case, and it concludes the proof of both theorems in the introduction.

References

[1] V. Bangert. On the existence of closed geodesics on two-spheres. *Internat. J. Math.*, 4(1):1–10, 1993.

[2] M. Berger. *Geometry revealed*. Springer, Heidelberg, 2010. A Jacob's ladder to modern higher geometry, Translated from the French by Lester Senechal.

[3] J.E. Borzellino, C.R. Jordan-Squire, G.C. Petrics, and D.M. Sullivan. Closed geodesics on orbifolds of revolution. *Houston J. Math.*, 33(4):1011–1025, 2007.

[4] T. Bröcker and T. tom Dieck. *Representations of compact Lie groups*, volume 98 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1985.

[5] J. Franks. Geodesics on $\mathbb{S}^2$ and periodic points of annulus homeomorphisms. *Invent. Math.*, 108(2):403–418, 1992.

[6] N. Hingston. Curve shortening, equivariant Morse theory, and closed geodesics on the 2-sphere. In *Differential geometry: Riemannian geometry (Los Angeles, CA, 1990)*, volume 54 of *Proc. Sympos. Pure Math.*, pages 423–429. Amer. Math. Soc., Providence, RI, 1993.

[7] N. Hingston. On the growth of the number of closed geodesics on the two-sphere. *Internat. Math. Res. Notices*, (9):253–262, 1993.

[8] A. Katok. Entropy and closed geodesics. *Ergodic Theory Dynam. Systems*, 2(3-4):339–365 (1983), 1982.

[9] A. Katok. Four applications of conformal equivalence to geometry and dynamics. *Ergodic Theory Dynam. Systems*, 8:139–152, 1988.

[10] S. Kobayashi. *Transformation groups in differential geometry*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1972 edition.

[11] G.A. Margulis. Certain applications of ergodic theory to the investigation of manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 3(4):89–90, 1969.

[12] A. Oancea. Morse theory, closed geodesics, and the homology of free loops spaces. In *Free loop spaces in geometry and topology*, volume 24 of *IRMA Lectures in Mathematics and Theoretical Physics*. European Mathematical Society, 2015.

[13] P. Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2006.

[14] A. Walfisz. *Weylsche Exponentialsummen in der neueren Zahlentheorie*. Mathematische Forschungsberichte, XV. VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.

Department of Mathematics, University of Oklahoma, Norman, Oklahoma, 73019
E-mail address: kennard@math.ou.edu

Department of Mathematics, SISSA International School for Advanced Studies, Via Bonomea, 265, 34136 Trieste, Italy
E-mail address: jrainone@sissa.it