J. DIFFERENTIAL GEOMETRY
39 (1994) 229–249

A SURFACE WITH POSITIVE CURVATURE AND POSITIVE TOPOLOGICAL ENTROPY

GERHARD KNIEPER & HOWARD WEISS

Abstract

We construct explicit examples of closed surfaces with positive curvature whose geodesic flow has positive topological entropy. It follows that these surfaces have infinitely many hyperbolic closed geodesics.

1. Introduction

We begin by posing the following general question:

Question. *Let M be a C^∞ closed orientable surface and g a smooth Riemannian metric with positive Gaussian curvature. Can the geodesic flow for g have a complicated dynamical behavior?*

The Gauss-Bonnet Theorem tells us that a positively curved surface must be topologically a sphere. The most common examples are the round sphere (and other surfaces of revolution) and the tri-axial ellipsoid. Both of these examples possess simple dynamics (i.e., their geodesic flows are not ergodic and they have zero entropies.) One might think that the simple topology of the sphere could be an obstruction for the geodesic flow of g to have complicated dynamics. This is not the case. Donnay [7] and Burns and Gerber [3] have constructed smooth (and real analytic) metrics on the sphere whose geodesic flows are Bernoulli. Katok [12] has shown that the simple topology is not an obstruction for a map to possess complicated dynamical behavior by constructing Bernoulli diffeomorphisms of the 2-disk.

Donnay, Burns, and Gerber construct their metrics by starting with a thrice punctured sphere and considering its complete hyperbolic metric. They then alter the metric far off into the cusps by cutting off the remainder of the cusps and glueing in reflecting caps. The geodesics leave the reflecting caps focused as they entered, and the cone family can be controlled in the caps. It is clear that these examples have "mostly" negative

Received February 11, 1991 and, in revised form, September 25, 1992. The work of the first author was supported by the SFB 170 in Göttingen and the work of the second author was partially supported by a National Science Foundation Postdoctoral Research Fellowship.

curvature, and that the negative curvature is the mechanism that causes the complicated dynamics. Although later examples by these authors require less negative curvature, *some* negative curvature is essential for their constructions.

The outstanding open question in this subject is whether such an example exists for a positively curved metric; i.e., does there exist a smooth Riemannian metric on S^2 with positive curvature whose geodesic flow is ergodic or has positive Liouville entropy? At present, this problem seems totally intractable. In this paper we consider the more modest question: Does there exist a smooth Riemannian metric on the sphere with positive curvature whose geodesic flow has positive topological entropy? We answer this question in the affirmative.

Theorem. *There exist "many" smooth Riemannian metrics on S^2 with positive Gaussian curvature whose geodesic flows have positive topological entropy.*

We conjecture that the *generic* metric on S^2 has positive topological entropy.

We prove our theorem by finding a *horseshoe* in the dynamics of the geodesic flow. A flow is said to possess an exponential growth rate of closed orbits if the number of closed orbits with primitive period $\leq T$ grows exponentially in T. The existence of a horseshoe guarantees the existence of infinitely many hyperbolic periodic orbits having an exponential growth rate. This gives the following corollary.

Corollary. *There exist "many" smooth Riemannian metrics on S^2 with positive Gaussian curvature having an exponential growth rate of hyperbolic closed geodesics.*

In particular, these metrics are the first known examples of positively curved metrics on S^2 having *infinitely many* hyperbolic closed geodesics. Our examples are obtained from small local perturbations of an ellipsoid E with distinct axes. The ellipsoid has the following remarkable property: E possesses two hyperbolic closed geodesics c and $-c$ (the same curve on E but with different orientations) such that any geodesic that is forwards asymptotic to c is backwards asymptotic to $-c$, and any geodesic that is forwards asymptotic to $-c$ is backwards asymptotic to c. By studying a suitable Poincaré section, this property implies that the Poincaré map for the flow possesses a double heteroclinic connection. The philosophy is that a double heteroclinic connection should be highly unstable and *both* connections can be broken and made transverse by a "generic" small perturbation. The subtlety is that we must perturb the Poincaré map (living in the unit tangent bundle of E) by perturbing the metric on E. This severely constrains the set of allowable perturbations. Also, *many* metric perturbations will affect *both* branches of the connection.

It is well known that a smooth map f that possesses transverse heteroclinic points contains a horseshoe, i.e., there exists a closed invariant subset on which f is topologically conjugate to a subshift of finite type. This implies that f has positive topological entropy [21]. We show that one only needs a two-sided crossing of the stable and unstable manifolds to obtain positive entropy. This is a new observation. Abramov's Theorem then implies that the geodesic flow has positive topological entropy.

Jacobi showed that the geodesic flow on the ellipsoid is a completely integrable, nondegenerate Hamiltonian system. Hence, the Kolmogorov-Arnold-Moser (KAM) Theorem tells us that our examples are not ergodic; the geodesic flows for our metrics exhibit complicated dynamics on a subset of the phase space TM.

Our method was inspired by the Poincaré method [19], [18, §403]. However, the Poincaré method is not applicable to autonomous systems. To extend this method to our construction, we must exploit the complete integrability of the geodesic flow on E. In short, we replace the unperturbed Hamiltonian in the Poincaré integral with the second integral for the geodesic flow on E.

Katok [11] has also constructed small Finsler perturbations of the round sphere with ergodic geodesic flows. The KAM Theorem does not apply in this case because the geodesic flow on the round sphere is a "degenerate" integrable system. However, these examples have zero topological entropy.

Donnay [8], using a simplified version of our methods, has constructed families of small local perturbations of the elliptical billiard having positive topological entropy.

Our examples are small perturbations of the ellipsoid by symmetric 2-forms having support in a small disk; hence they are clearly not real analytic. However, we can approximate our metrics which have positive entropy arbitrarily well in the C^∞ topology using real analytic metrics. If one combines results of Katok, Newhouse, and Yomdin [13], [14], [16], [26], one obtains that the topological entropy of C^∞ flows on closed three-dimensional manifolds is continuous. Hence, we obtain real analytic examples.

Theorem. *There exist "many" real analytic Riemannian metrics on S^2 with positive Gaussian curvature whose geodesic flows have positive topological entropy.*

2. Geometry of the ellipsoid

The geometry of the ellipsoid can be neatly described using the *elliptic coordinates* of Jacobi. We quickly review the most germane geometric facts. Our review is based on the exposition in [15]. Other references include [2], [10], [20], [22], and [23].

We consider the two-dimensional ellipsoid $E = \{x_0^2/a_0 + x_1^2/a_1 + x_2^2/a_2 = 1\}$, where we assume that $0 < a_0 < a_1 < a_2$. The intersections of E with the three coordinate planes yield simple closed geodesics which are called the *basic closed geodesics*. E has four *umbilic points*

$$(x_0, x_1, x_2) = (\pm\sqrt{a_0}\sqrt{a_1 - a_0}/\sqrt{a_2 - a_0}, 0, \pm\sqrt{a_2}\sqrt{a_2 - a_1}/\sqrt{a_2 - a_0}),$$

which are located on the middle length basic closed geodesic. These points play a crucial role in the study of the geometry of E.

For $\rho \notin \{a_0, a_1, a_2\}$, consider the function $A_\rho : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$A_\rho(x_0, x_1, x_2) = \frac{x_0^2}{a_0 - \rho} + \frac{x_1^2}{a_1 - \rho} + \frac{x_2^2}{a_2 - \rho}.$$

It is an easy exercise in analytic geometry to show that through each point p of E outside the coordinate planes, there passes exactly one one-sheeted hyperboloid $\{A_{u_1} = 1\}$ and one two-sheeted hyperboloid $\{A_{u_2} = 1\}$, where $(u_1, u_2) \in (a_0, a_1) \times (a_1, a_2)$. We may take $u = (u_1, u_2)$ as local coordinates and extend them to local coordinates on all of E. In these coordinates, the metric tensor $ds^2 = (-u_1 + u_2)(U_1 du_1^2 + U_2 du_2^2)$ with $U_i = U_i(u_i) = (-1)^i u_i / f(u_i)$ and $f(u_i) = 4(a_0 - u_i)(a_1 - u_i)(a_2 - u_i)$. Hence the Hamiltonian for the geodesic flow is given by

$$H(u, \dot{u}) = (-u_1 + u_2)(U_1 \dot{u}_1^2 + U_2 \dot{u}_2^2).$$

The Gaussian curvature of E is given by

$$K(u_1, u_2) = a_0 a_1 a_2 / u_1^2 u_2^2,$$

where $(u_1, u_2) \in [a_0, a_1] \times [a_1, a_2]$. The geodesics on E (that do not pass through the umbilics) are characterized by

$$\frac{\sqrt{U_1}}{\sqrt{-u_1 + \gamma}} \dot{u}_1 \pm \frac{\sqrt{U_2}}{\sqrt{u_2 - \gamma}} \dot{u}_2 = 0, \qquad \gamma \in (a_0, a_1) \text{ or } (a_1, a_2),$$

together with the condition $H(u, \dot{u}) = 1$. From this characterization of geodesics, we can determine the second integral for the geodesic flow:

$$F(u, \dot{u}) = (-u_1 + u_2)(u_2 U_1 \dot{u}_1^2 + u_1 U_2 \dot{u}_2^2),$$

or more geometrically,

$$F(v) = u_1(\pi \circ v)\sin^2 \mu(v) + u_2(\pi \circ v)\cos^2 \mu(v),$$

where $v \in SE$ (the unit tangent bundle to E), $\pi\colon TM \to M$ is the canonical projection, and $\mu(v)$ denotes the angle between v and the u_1-parameter line through $\pi \circ v$. One can easily verify that H and F are in involution, i.e., $\{H, F\} = 0$, and that dH and dF are everywhere linearly independent.

Hamilton's equations imply that the projections to E of the Hamiltonian flows associated to F and H satisfy the differential equations:

$$\frac{du_1^F}{dt} = \frac{dF}{d\dot{u}_1} = (-u_1 + u_2)(2u_2 U_1 \dot{u}_1),$$
$$\frac{du_2^F}{dt} = \frac{dF}{d\dot{u}_2} = (-u_1 + u_2)(2u_1 U_1 \dot{u}_2),$$
$$\frac{du_1^H}{dt} = \frac{dH}{d\dot{u}_1} = (-u_1 + u_2)(U_1 \dot{u}_1),$$
$$\frac{du_2^H}{dt} = \frac{dH}{d\dot{u}_2} = (-u_1 + u_2)(U_1 \dot{u}_2).$$

It is clear that these projected flows are independent if $u_1 \neq u_2$ and $\dot{u}_1 \neq 0 \neq \dot{u}_2$.

For $\gamma \in (a_1, a_2)$, the (geodesic) flow invariant set $\{F = \gamma\}$ is a union of two two-dimensional tori $T_\gamma^\pm$ whose flow lines project to geodesics that monotonically wind around the x_2-axis and oscillate between the two u_1-parameter lines $\{u_2 = \gamma\}$. Similarly, for $\gamma \in (a_0, a_1)$, the flow invariant set $\{F = \gamma\}$ is a union of two two-dimensional tori $T_\gamma^\pm$ whose flow lines project to geodesics that monotonically wind around the x_0-axis and oscillate between the two u_2-parameter lines $\{u_1 = \gamma\}$. The coordinates on $T_\gamma^\pm$ that linearize the flow are the integral curves of the vector fields generated by H and $F - 2\gamma H$. It follows from this analysis that the closed geodesics are dense in SE and that the shortest and longest basic closed geodesics are elliptic.

The behavior of the flow near the middle length basic closed geodesic c is more interesting. The closed geodesic c is hyperbolic (see proof below). The flow invariant set $SE \cap \{F = a_1\}$ corresponds to geodesics on E that pass through an umbilic point. The geodesics c and $-c$ are the only closed geodesics contained in this set (see Appendix).

Every other geodesic with tangent vectors in $SE \cap \{F = a_1\}$ passes through one of the two pairs $\{q, q'\}$ or $\{r, r'\}$ of diametrical umbilics

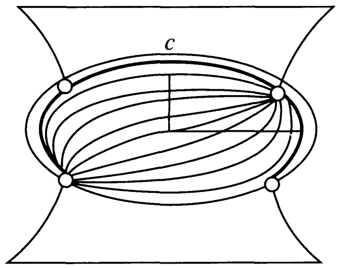

FIGURE 1. GEODESICS SPIRALING INTO c.

time and again at fixed intervals. This interval T is the same for all flow lines. It follows from this analysis that every geodesic that is forwards asymptotic to c is backwards asymptotic to $-c$ and passes through the pair $\{q, q'\}$ of umbilics infinitely many times (see Figure 1). One can give an easy dynamics proof of this proposition by considering the geodesic flow for time $2T$ restricted to the unit tangent vectors having a footpoint on an umbilic, say q. This induced map is a circle diffeomorphism having rotation number 0 and two hyperbolic fixed points. The proposition then follows from the standard results in the theory of circle diffeomorphisms [6].

By this analysis, the weak stable manifold of c (the set of vectors whose geodesics are forward asymptotic to c) coincides with the weak unstable manifold for $-c$ (the set of vectors whose geodesics are backward asymptotic to $-c$) and forms a double branched cover of E, branched over c.

We now show that the middle length basic closed geodesic c on E is hyperbolic (see [15]). This is a generalization of the classical fact that any billiard path in an elliptic billiard that passes through a focus, is both forwards and backwards asymptotic to the major axis [5].

We parametrize c such that c starts at $c(0) = (\sqrt{a_0}, 0, 0)$, $\dot{c}$ points into the half-space $\{x_2 > 0\}$, and c has length l. Let $0 < t_1 < t_2 < t_3 < t_4 < l$ be the parameter values where c passes through the four umbilics.

Theorem 2.1. *The middle length basic closed geodesic* c *on E is hyperbolic.*

Proof. We will show that $E^s = \{$orthogonal Jacobi fields J along $c \mid J(t_1) = 0$ and $J'(t_1) > 0\}$ generates the stable eigenspace of the Poincaré map with eigenvalue $0 < a < 1$. From the geometry of the ellipsoid we know that $J(t_1) = 0$ implies $J(t_3) = 0 = J(t_1 + l)$. This implies that $t \mapsto J(t + l) \in E^s$, and since the dimension of $E^s = 1$, there exists $a \in \mathbb{R}$ such that $J(t + l) = a \cdot J(t)$. We will show that $0 < a < 1$. Then

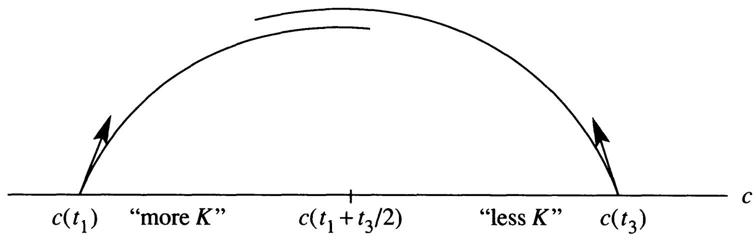

FIGURE 2. JACOBI FIELDS J_1 AND J_2.

$0 < J'(t_1 + l) = a \cdot J'(t_1) < J'(t_1)$ and $J(t_1 + l) = J(t_1) = 0$. This clearly implies the theorem.

Lemma 2.2. *If* $J \in E^s$, *then* $-J'(t_3) < J'(t_1)$.

Consider the interval $[t_1, t_3]$. By studying the simple expression for the Gaussian curvature given above, one can easily show that $K(t_1 + s) > K(t_3 - s)$ for $0 < s < (t_3 + t_1)/2$. The idea is that the curvature formula implies that $K(t)$ has the property that if t_2 is the time at which c passes through the "intermediate" umbilic point, then $K(t_1 + s) = K(t_2 - s) \geq K(t_1)$ for $0 \leq s \leq t_2 - t_1$ and $K(t_2 + s) = K(t_3 - s) \leq K(t_1)$ for $0 \leq s \leq t_3 - t_2$.

Now consider two arbitrary Jacobi fields J_1 and J_2 along c such that $J_1(t_1) = J_2(t_3) = 0$ and $J_1'(t_1) \leq -J_2'(t_3)$. Since there are no conjugate points between t_1 and t_3, the Sturm Comparison Theorem implies that $0 < J_1((t_1 + t_3)/2) < J_2((t_1 + t_3)/2)$. Now consider our original Jacobi field J. Call it J_1 when restricted to $[t_1, (t_1 + t_3)/2]$, and J_2 on $[(t_1 + t_3)/2, t_3]$. The previous argument shows that $-J'(t_3) < J'(t_1)$, for otherwise, J_1 and J_2 would not agree at the midpoint of the interval (see Figure 2). q.e.d.

Since $J'(t_1) > 0$, and the only zeros of J occur at t_1 and t_3 modulo l, we have that $J'(t_1 + l) > 0$. If we apply the argument in the proof of the Lemma to the interval $[t_3, t_1 + l]$, we obtain that $0 < J'(t_1 + l) < -J'(t_3)$. Combining this statement with the lemma, we obtain that $0 < J'(t_1 + l) < -J'(t_3) < J'(t_1)$.

3. Splitting of stable and unstable manifolds

Let M be a C^∞ n-dimensional compact Riemannian manifold. Hamiltonian flows are usually flows on T^*M (the cotangent bundle to M). However, we are mostly interested in geodesic flows which are more naturally thought of as flows on TM (the tangent bundle to M). Using the Riemannian metric, we may transfer the natural symplectic structure on T^*M to TM and consider Hamiltonian flows on TM. The Hamiltonian function $H : TM \to \mathbb{R}$ for geodesic flows is given by the Riemannian metric, i.e., $H(v) = g(v, v)$. Let $\pi : TM \to M$ be the canonical projection.

Assume that our system has two closed orbits $\alpha(t)$ and $\omega(t)$ (not necessarily distinct) and a biasymptotic orbit $x(t)$. We will assume that $x(t)$ is backwards asymptotic to $\alpha(t)$ and forwards asymptotic to $\omega(t)$. We consider Hamiltonian perturbations of our flow $H_\epsilon = H_0 + \epsilon H_1$ such that the support of H_1 is bounded away from $\alpha(t)$ and $\omega(t)$. This insures that α and ω will not be affected by the perturbation. Under these assumptions we can prove the following theorem.

Theorem 3.1. *Let $\phi_0^t : TM \to TM$ be the Hamiltonian flow for the Hamiltonian H_0. Let $H_\epsilon = H_0 + \epsilon H_1$ be a perturbation such that the support of H_1 is bounded away from $\alpha(t)$ and $\omega(t)$. Let ϕ_ϵ^t be the associated Hamiltonian flows on $H_\epsilon = 1$. Assume that for ϵ sufficiently small the biasymptotic orbit $x(t)$ persists under the perturbation H_ϵ, i.e., there exists a family of orbits $x_\epsilon(t)$ such that $x_0(t) = x(t)$, $d(x_\epsilon(t), \alpha(t)) \to 0$ for $t \to -\infty$, $d(x_\epsilon(t), \omega(t)) \to 0$ for $t \to \infty$, and $d(x_\epsilon(t), x(t)) = O(\epsilon)$ uniformly in t. Furthermore, suppose that the flow associated to H_0 possesses a smooth integral of motion $F : TM \to \mathbb{R}$ which is in involution with H_0, i.e., $\{H_0, F\} = 0$. Then*

$$\int_{-\infty}^{+\infty} \{F, H_1\}(x(t))\, dt = 0.$$

Remarks. (1) Since the support of H_1 is bounded away from α and ω, the integral always exists and is finite.

(2) If $F = H_0$, this integral is called the Poincaré integral (or the Poincaré-Melnikov integral) and vanishes since $\frac{d}{dt}H_1(x(t))\,dt = \{H_1, H_0\}(x(t))$.

(3) For the geodesic flow on the ellipsoid E, $W^{ws}(c) = W^{wu}(-c)$ and $W^{ws}(-c) = W^{wu}(c)$ where c and $-c$ are the hyperbolic closed geodesics of "middle length", and $W^{ws}(c)$ and $W^{wu}(c)$ denote the weak stable and weak unstable manifolds for c.

Proof. Since F is an integral of motion of the unperturbed system H_0, there are constants c_1, c_2, and c_3 such that $F(\alpha(t)) = c_1$, $F(\omega(t)) = c_2$, and $F(x(t)) = c_3$ for all $t \in \mathbb{R}$. Since $x(t)$ is biasymptotic, the continuity of F implies that the constants coincide.

Lemma 3.2. *Under the conditions of the theorem we have*

$$\int_{-\infty}^{+\infty} \frac{d}{dt} F(x_\epsilon(t))\, dt = 0.$$

Proof of Lemma.

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{d}{dt} F(x_\epsilon(t))\, dt &= \lim_{T\to\infty} \int_{-T}^{+T} \frac{d}{dt} F(x_\epsilon(t))\, dt \\ &= \lim_{T\to\infty} (F(x_\epsilon(T)) - F(x_\epsilon(-T))) = 0,\end{aligned}$$

since $x_\epsilon(t) \to \omega$ as t $\to \infty$, $x_\epsilon(t) \to \alpha$ as $t \to -\infty$, and $F(\alpha) = F(\omega)$. q.e.d.

We continue with the proof of our Theorem 3.1. Using the Lemma and the definition of Poisson bracket we obtain

$$\begin{aligned}0 &= \int_{-\infty}^{+\infty} \frac{d}{dt} F(x_\epsilon(t))\, dt = \int_{-\infty}^{+\infty} \{F, H_\epsilon\}(x_\epsilon(t))\, dt \\ &= \int_{-\infty}^{+\infty} \{F, H_0 + \epsilon H_1\}(x_\epsilon(t))\, dt \\ &= \int_{-\infty}^{+\infty} \{F, H_0\}(x_\epsilon(t))\, dt + \epsilon \int_{-\infty}^{+\infty} \{F, H_1\}(x_\epsilon(t))\, dt.\end{aligned}$$

The first integral above vanishes since F and H_0 are in involution. We are left with

$$\int_{-\infty}^{+\infty} \{F, H_1\}(x_\epsilon(t))\, dt = 0, \quad \text{for sufficiently small } \epsilon.$$

Since H_1 has support bounded away from $\alpha(t)$ and $\omega(t)$, this integral is between *finite* limits. We can take the limit inside the integral and since $d(x_\epsilon(t), x(t)) = O(\epsilon)$, we obtain

$$\int_{-\infty}^{+\infty} \{F, H_1\}(x(t))dt = 0.$$

4. Splitting of asymptotic manifolds

Given a hyperbolic closed orbit γ of a flow, the weak stable (weak unstable) manifold consists of the set of points that are forwards (backwards) asymptotic to γ. The Stable Manifold Theorem [9] tells us that the weak stable and unstable manifolds for γ are smooth surfaces that vary smoothly when the flow is smoothly perturbed. Let α and ω be two hyperbolic closed geodesics, and let $W_\epsilon^{wu}(\alpha)$ denote the weak unstable manifold of α and $W_\epsilon^{ws}(\omega)$ the weak stable manifold of ω with respect to the system $H_\epsilon = H_0 + \epsilon H_1$. The following Corollary is an immediate consequence of Theorem 3.1.

Corollary 4.1. *Assume that* $W_0^{wu}(\alpha) = W_0^{ws}(\omega)$. *If there exists a biasymptotic orbit* $x(t)$ *of* H_0 *such that* $x(t) \in W_0^{wu}(\alpha) = W_0^{ws}(\omega)$, *and if there exists a second integral of motion* $F : TM \to \mathbb{R}$ *that is independent of* H_0 *and in involution with* H_0 *such that*

$$\int_{-\infty}^{\infty} \left.\frac{d}{ds}\right|_{s=0} H_1(\phi_F^s(x(t)))\,dt \neq 0,$$

then $W_\epsilon^{wu}(\alpha) \neq W_\epsilon^{ws}(\omega)$ *for* ϵ *sufficiently small.*

Proof of Corollary. Assume that $W_\epsilon^{wu}(\alpha) = W_\epsilon^{ws}(\omega)$ for sufficiently small ϵ. By the continuous dependency of stable and unstable manifolds on the perturbation parameter [9], we can find a family of orbits $x_\epsilon \in W_\epsilon^{wu}(\alpha) = W_\epsilon^{ws}(\omega)$ such that $d(x_\epsilon(t), x(t)) = O(\epsilon)$. Since $\left.\frac{d}{ds}\right|_{s=0} H_1(\phi_F^s(x(t)))\,dt = \{H_1, F\}(x(t))$, an application of Theorem 3.1 yields the corollary.

Warning. The condition that $W_\epsilon^{wu}(\alpha) \neq W_\epsilon^{ws}(\omega)$ for ϵ sufficiently small does not necessarily imply that the two manifolds intersect transversely; they may be disjoint or they may intersect nontransversely with a one-sided or two-sided crossing.

5. A local conformal perturbation of the ellipsoid which produces positive topological entropy

Let $x(t)$ and $y(t)$ be biasymptotic orbits such that there are constants $\sigma_+, \tau_+, \sigma_-, \tau_-$ such that $x(t) \to \dot{c}(t+\sigma_+)$ and $y(t) \to \dot{c}(t+\tau_+)$ as $t \to \infty$ and $x(t) \to -\dot{c}(t+\sigma_-)$ and $y(t) \to -\dot{c}(t+\tau_-)$ as $t \to -\infty$. We also choose the parametrizations of $x(t)$ and $y(t)$ so that $\pi \circ x(0)$ and $\pi \circ y(0)$ do not lie on c and are distinct points. Let N, N' be the respective neighborhoods of $\pi \circ x(0), \pi \circ y(0)$ such that $N \cap N' = \varnothing$, $N \cap c = \varnothing$, $N' \cap c = \varnothing$, $(\pi \circ x(t)) \cap N' = \varnothing$, $(\pi \circ y(t)) \cap N = \varnothing$, $\pi \circ x(t)$ does not reenter N, and $\pi \circ y(t)$ does not reenter N'.

We want to use the map $G_N(s, t) = \pi(\phi_H^s x(t)) = \pi(\phi_F^s \circ \phi_H^t(x(0)))$ as a coordinate system in a neighborhood of $\pi \circ x(0)$. It is clear from the remarks in §2 that this will be possible provided $u_1 \neq u_2$ and $\dot{u}_1 \neq 0 \neq \dot{u}_2$ at $x(0)$. The first condition holds because $\pi \circ x(0)$ is not an umbilic point; the second condition holds because $x(0)$ is tangent to a geodesic that passes through the umbilic points. Thus $G_N : [-\delta, \delta] \times [-\epsilon, \epsilon] \to N$ will provide local coordinates for N for sufficiently small ϵ and δ.

We now define a function $P_N : [-\delta, \delta] \times [-\epsilon, \epsilon] \to \mathbb{R}$ by

$$P_N(s, t) = s \cdot t^2 \cdot \chi_{[-\delta, \delta]}(s) \cdot \kappa_{[-\epsilon, \epsilon]}(t),$$

where $\chi_{[-\delta, \delta]}$, $\kappa_{[-\epsilon, \epsilon]}$ are smooth positive functions supported in $[-\delta, \delta]$ and $[-\epsilon, \epsilon]$ respectively with $\chi(0) = \kappa(0) = 1$. Let

$$H_{1,N}(v) := \begin{cases} P_N(G_N^{-1} \circ \pi(v)) \cdot H_0(v), & \text{if } \pi(v) \in \text{image}(G_N) \subset N, \\ 0, & \text{otherwise.} \end{cases}$$

If $H_0 = g$, we obtain a conformal perturbation

$$g_{\epsilon, N} = H_0 + \epsilon H_{1,N} = H_0 \cdot (1 + \epsilon P_N(G_N^{-1} \circ \pi)) = (1 + \epsilon P_N) \cdot g.$$

For this perturbation we have

$$\begin{aligned} \left.\frac{d}{ds}\right|_{s=0} H_{1,N}(\phi_F^s x(t)) &= \left.\frac{d}{ds}\right|_{s=0} P_N(G_N^{-1} \circ \pi(\phi_F^s x(t)) \\ &= \left.\frac{d}{ds}\right|_{s=0} P_N(s, t) = t^2 \cdot \kappa_{[-\epsilon, \epsilon]}(t). \end{aligned}$$

This construction yields

$$\int_{-\infty}^{\infty} \left.\frac{d}{ds}\right|_{s=0} H_{1,N}(\phi_F^s x(t))\, dt = \int_{-\infty}^{\infty} t^2 \kappa_{[-\epsilon, \epsilon]}(t)\, dt > 0.$$

We repeat the preceding construction in N' to obtain a local coordinate system $G_{N'}$, a conformal factor $P_{N'}$, a Riemannian metric $g_{\epsilon, N'}$ and a Hamiltonian function $H_{1,N'}$ such that

$$\int_{-\infty}^{\infty} \left.\frac{d}{ds}\right|_{s=0} H_{1,N'}\left(\phi_F^s(y(t))\right) dt > 0.$$

We combine these two local conformal perturbations by defining $H_1 = H_{1,N} + H_{1,N'}$, i.e.,

$$H_1(v) := \begin{cases} P_N(G_N^{-1} \circ \pi(v)) \cdot H_0(v), & \text{if } \pi(v) \in \text{image}(G_N) \subset N, \\ P_{N'}(G_{N'}^{-1} \circ \pi(v)) \cdot H_0(v), & \text{if } \pi(v) \in \text{image}(G_{N'}) \subset N', \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\begin{aligned} g_\epsilon &= H_0 + \epsilon H_1 = H_0 \cdot (1 + \epsilon P_N(G_N^{-1} \circ \pi) + \epsilon P_{N'}(G_{N'}^{-1} \circ \pi)) \\ &= (1 + \epsilon P_N + \epsilon P_{N'}) \cdot g. \end{aligned}$$

Since $c \cap (N \cup N') = \varnothing$, c is not affected by the perturbation. Thus by Corollary 4.1, $W_\epsilon^{ws}(\dot{c}(t)) \neq W_\epsilon^{wu}(-\dot{c}(t))$ and $W_\epsilon^{ws}(-\dot{c}(t)) \neq W_\epsilon^{wu}(\dot{c}(t))$

for ϵ sufficiently small. It will follow from the results in §7 that the geodesic flow for g_ϵ has positive topological entropy for $\epsilon > 0$ sufficiently small.

6. Local genericity of positive topological entropy

In this section we show that the geodesic flow for a generic metric perturbation of $H_0 = g$ having support contained in $N \cup N'$ has positive topological entropy.

Let N and N' be as in §5. Let $\mathfrak{Sym}_{2,N\cup N'}(M)$ denote the space of symmetric 2-tensors with support in $N \cup N'$. Then the function K : $\mathfrak{Sym}_{2,N\cup N'}(M) \to \mathbb{R}^2$ defined by

$$K\colon \xi \to \left(\int_{-\infty}^{+\infty} \frac{d}{ds}\bigg|_{s=0} \xi(\phi_F^s(x(t)), \phi_F^s(x(t)))\, dt, \right.$$
$$\left. \int_{-\infty}^{+\infty} \frac{d}{ds}\bigg|_{s=0} \xi(\phi_F^s(y(t)), \phi_F^s(y(t)))\, dt \right)$$

is clearly linear and locally bounded. In §5, we showed that the function K does not vanish identically by explicitly constructing a metric perturbation H_1 of H_0 such that both integrals do not vanish.

Choose $\xi \in \mathfrak{Sym}_{2,N\cup N'}(M)$ and consider the integral $K(\xi)$. If this integral does not vanish, we need not do anything further. If it does vanish, then the local continuity of K implies that we may slightly perturb ξ such that the integral does not vanish. In fact, an *open, dense subset* of $\mathfrak{Sym}_{2,N\cup N'}(M)$ will have nonvanishing integrals, and the results in §7 will imply that the corresponding geodesic flows will have positive topological entropy.

7. The two-sided intersection of weak stable and weak unstable manifolds

Consider the set of all the unit tangent vectors having footpoint on the long basic closed geodesic l, and remove the two tangent vectors pointing along l. What remains has two connected components—both annuli. Let Σ denote the Poincaré section consisting of one of these annuli. The corresponding Poincaré map preserves a natural smooth 2-form and possesses a pair of nontransverse heteroclinic points v and $-v$, where v is tangent to c. In §5 we constructed an explicit local metric perturbation $\{g_\epsilon\}$ of E. The corresponding geodesic flows induce Poincaré maps $\{P_\epsilon\}$ on Σ, which preserve smooth 2-forms.

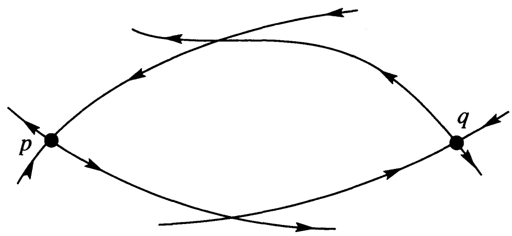

FIGURE 3. HETEROCLINIC POINTS WITH 2-SIDED CROSSING.

If a diffeomorphism f has a hyperbolic fixed point p, the stable manifold $W^s(p)$ of p consists of the set of points that are forward asymptotic to p and the unstable manifold $W^u(p)$ of p consists the set of points that are backwards asymptotic to p.

It follows from the nonvanishing of the two integrals for our explicit metric perturbation in §5 that $W_\epsilon^s(v) \neq W_\epsilon^u(-v)$ and $W_\epsilon^s(-v) \neq W_\epsilon^u(v)$ for the perturbed Poincaré maps $\{P_\epsilon\}$. To show that the geodesic flow for the perturbed metric g_ϵ has positive topological entropy, we will show that $W_\epsilon^s(v)$ has a two-sided intersection with $W_\epsilon^u(-v)$, and $W_\epsilon^s(-v)$ has a two-sided intersection with $W_\epsilon^u(v)$.

We required our metric perturbation to be supported in N and N' because for the metric deformation $g_{\epsilon,N}$ supported in N, one can only conclude that $W_\epsilon^s(v)$ has a two-sided intersection with $W_\epsilon^u(-v)$ for the corresponding Poincaré maps. It is possible, although highly improbable, that $W_\epsilon^s(-v) = W_\epsilon^u(v)$. In this case, it is not clear whether the Poincaré maps have positive topological entropy. We elected to eliminate this possibility by considering a metric perturbation supported in N and N', which was engineered to split both connections and produce two-sided crossings for the Poincaré map. See Figure 3.

We require the following simple lemma.

Lemma 7.1. *Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a smooth diffeomorphism that preserves a smooth 2-form ω and has a double heteroclinic connection, i.e., there exist hyperbolic fixed points p and q such that $W^s(p) = W^u(q)$ and $W^s(q) = W^u(p)$. Let N denote a small disk such that $N \cap \{p, q\} = \varnothing$, $N \cap W^s(p) \neq \varnothing$, and $N \cap W^s(q) = \varnothing$. Let f_ϵ, $f_0 = f$ be a smooth deformation of f that is supported in N and which preserves smooth 2-forms ω_ϵ that vary smoothly in ϵ. If $W_\epsilon^s(p) \neq W_\epsilon^u(q)$, then $W_\epsilon^s(p)$ and $W_\epsilon^u(q)$ intersect with a two-sided crossing for ϵ sufficiently small. See Figure* 4 (*next page*).

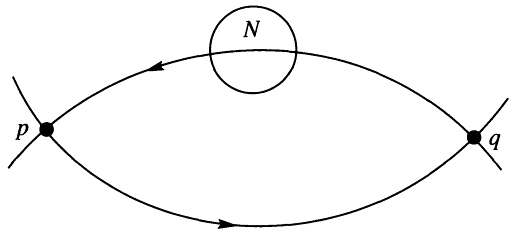

FIGURE 4. HYPOTHESIS OF LEMMA 7.1.

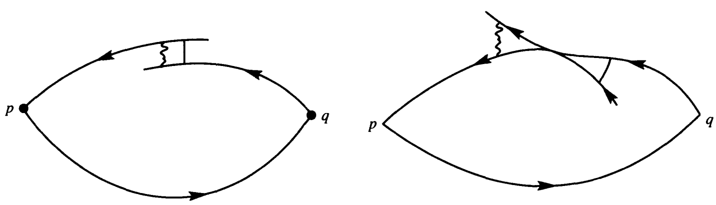

FIGURE 5. FORBIDDEN CONFIGURATIONS.

Proof. In [18, §308], Poincaré proved a related result for homoclinic points. He proved that if f is an area-preserving surface diffeomorphism and p is a hyperbolic fixed point such that $W^s(p) = W^u(p)$, and if one considers a smooth deformation f_ϵ of f by area-preserving diffeomorphisms such that the stable and unstable manifolds of the perturbed hyperbolic fixed point p_ϵ do not coincide, then the perturbed stable and unstable manifolds for p intersect with a two-sided crossing. Poincaré considered the possible ways that the perturbed stable and unstable manifolds might not intersect and then showed that each case leads to a violation of the area-preserving hypotheses.

The proof of our Lemma for heteroclinic points is similar to Poincaré's argument. We illustrate the two main cases of the Lemma in Figure 5—the case when $W^s_\epsilon(p) \cap W^u_\epsilon(q) = \emptyset$, and the case $W^s_\epsilon(p) \cap W^u_\epsilon(q)$ with *only* one-sided crossings. Both cases violate the area preserving property of the perturbed Poincaré maps. The wiggly arc represents the image under f of the straight arc. Since N and $W^s(q)$ are disjoint, $W^s(q)$ does not feel the perturbation, and hence $W^s_\epsilon(q) = W^u_\epsilon(p)$.

The following proposition will ensure that the Poincaré maps for the local metric deformation will have heteroclinic points with two-sided crossings on both branches.

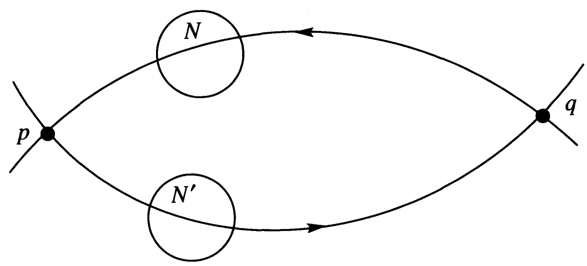

FIGURE 6. HYPOTHESIS OF PROPOSITION 7.2.

Proposition 7.2. *Let $f: \mathbb{R}^2 \to \mathbb{R}^2$ be a smooth diffeomorphism that preserves a smooth 2-form ω and has a double heteroclinic connection, i.e., there exist hyperbolic fixed points p and q such that $W^s(p) = W^u(q)$ and $W^s(q) = W^u(p)$. Let N and N' denote disjoint small disks such that $(N \cup N') \cap \{p, q\} = \varnothing$, $N \cap W^s(p) \neq \varnothing$, $N \cap W^s(q) = \varnothing$, $N' \cap W^s(q) \neq \varnothing$, and $N' \cap W^s(p) = \varnothing$. Let f_ϵ, $f_0 = f$, be a smooth deformation of f that is supported in $N \cup N'$ and preserves smooth 2-forms ω_ϵ that vary smoothly in ϵ. Suppose that $W^s_\epsilon(p) \neq W^u_\epsilon(q)$ and $W^s_\epsilon(q) \neq W^u_\epsilon(p)$ for ϵ sufficiently small. Then $W^s_\epsilon(p) \cap W^u_\epsilon(q) \neq \varnothing$, $W^s_\epsilon(q) \cap W^u_\epsilon(p) \neq \varnothing$, and both intersections have two-sided crossings for small positive ϵ.*

Proof. We may assume that $f_\epsilon(x) = f \circ g_\epsilon \circ h_\epsilon$ where support$(g_\epsilon) \subset N$ and support$(h_\epsilon) \subset N'$. Since $N \cap N' = \varnothing$, $g_\epsilon \circ h_\epsilon = h_\epsilon \circ g_\epsilon$. Consider the deformation $f \circ g_\epsilon$ supported in N. It follows from the hypothesis that for these maps, the stable manifold of q coincides with the unstable manifold of p, and the stable manifold at p does not coincide with the unstable manifold at q, for small positive ϵ. By Lemma 7.1, these manifolds must intersect with a two-sided crossing. Similarly, by reversing the roles of p and q we obtain that for the deformation $f \circ h_\epsilon$ that the stable manifold of p coincides with the unstable manifold at q and that the stable manifold of q intersects the unstable manifold at p with a two-sided crossing for small positive ϵ.

Since support$(h_\epsilon) \subset N'$, adding this perturbation to $f \circ g_\epsilon$ will not influence the two-sided crossing of the stable manifold at p and the unstable manifold at q. This implies that for the deformation $f \circ g_\epsilon \circ h_\epsilon$, $W^s_\epsilon(p)$ has a two-sided crossing with $W^u_\epsilon(q)$ for small positive ϵ. Similarly, we obtain that for the deformation $f \circ h_\epsilon \circ g_\epsilon$, $W^s_\epsilon(q)$ has a two-sided crossing with $W^u_\epsilon(p)$ for small positive ϵ.

Remarks. (1) These results in this chapter are also true for diffeomorphisms of an annulus, and the proofs are identical. We apply Proposition 7.2 in the case of an annulus.

(2) Proposition 7.2 is false if the perturbation is not supported away from p and q, and it is also false if the perturbation is large or non-area-preserving [25]. In addition, this Proposition is false in higher dimensions if we consider smooth perturbations of a symplectomorphism through symplectomorphisms.

8. Positive topological entropy

It is well known [21] that if $f: \mathbb{R}^2 \to \mathbb{R}^2$ is a diffeomorphism having hyperbolic fixed points p and p', such that $q \in W^s(p) \cap W^u(p')$ and $q' \in W^s(p') \cap W^u(p)$, where the intersections at q and q' are transverse, then f has positive topological entropy. One finds a horseshoe (a closed invariant Cantor set in M on which f is topologically equivalent to a subshift of finite type) in the dynamics of f. Since a subshift of finite type has positive topological entropy [24] and topological entropy is a topological invariant [24], this implies that f has positive topological entropy.

Remark. Suppose $f: \mathbb{R}^2 \to \mathbb{R}^2$ is a diffeomorphism having hyperbolic fixed points p and p', such that $W^s(p) = W^u(p')$ and $q' \in W^s(p') \cap W^u(p)$ where the intersection at q is transverse. It is not known whether f has positive topological entropy. Hence one can not conclude that the perturbed geodesic flows have positive topological entropy if the corresponding Poincaré maps satisfy the hypothesis in the Proposition in §6.

For heteroclinic points with two-sided crossings we show the following result, whose detailed proof appears in [25]:

Theorem 8.1. *Let $f: \mathbb{R}^2 \to \mathbb{R}^2$ be a diffeomorphism having hyperbolic fixed points p and p', such that $q \in W^s(p) \cap W^u(p')$ and $q' \in W^s(p') \cap W^u(p)$, and the intersections at q and q' are two-sided, then f has positive topological entropy.*

In the case when the stable and unstable manifolds have order r contact at a heteroclinic point $(1 \leq r < \infty)$ Conley [4] (see also [19]) showed that there exists a *transverse* heteroclinic point in every neighborhood of the nontransverse heteroclinic point. By the preceding remarks, this implies that f has positive topological entropy.

The case of infinite order contact is more delicate. It seems quite difficult to show directly the existence of a transverse heteroclinic point near an infinitely flat nontransverse heteroclinic point. However, it is straightforward to prove directly that f has positive topological entropy. One can then apply a result of Katok [13] which states that if a surface diffeomorphism has positive topological entropy, then there exists a horseshoe that carries most of the entropy.

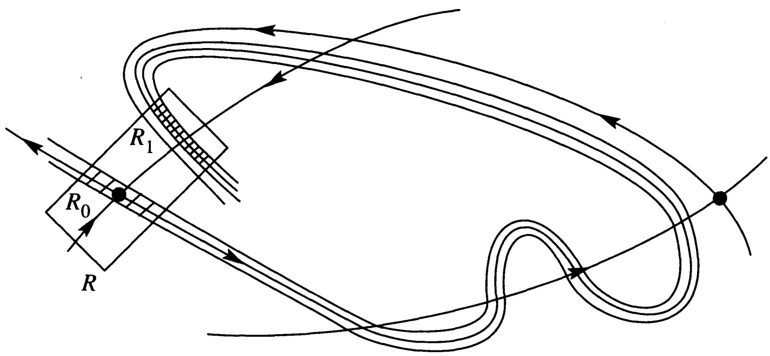

FIGURE 7. COMPONENTS OF $f^N(R) \cap R$.

To prove that f has positive topological entropy, one needs only to show that f has a topological factor that is a subshift of finite type, i.e., there exists a continuous surjection between an invariant subset of M and a subshift of finite type. The hard part of constructing the horseshoe in the transverse case is proving the injectivity of this map, and it requires the use of Palis' λ-Lemma, which is not applicable in our situation. However, injectivity of the map is not required to ensure that f has positive entropy. The idea of the proof is to choose a rectangle R to obtain a picture as in Figure 7, where R_1 and R_2 are two *good* components of $f^N(R) \cap R$, where f^N denotes some sufficiently large iterate of f. We use R_1 and R_2 to code our map and we show the continuity and surjectivity of the coding map by imitating the proof in the transverse case.

Remark. Theorem 8.1 also applies to diffeomorphisms of an annulus with the identical proof. It is in this form that we apply the theorem.

Proof of main theorem. Our goal is to show that the geodesic flow for the local conformal metric perturbation g_ϵ constructed in §5 has positive topological entropy. We observed in §7 that $W_\epsilon^s(v) \neq W_\epsilon^u(-v)$ and $W_\epsilon^s(-v) \neq W_\epsilon^u(v)$ for the perturbed Poincaré maps $\{P_\epsilon\}$. It follows immediately from Proposition 7.2 that $W_\epsilon^s(v) \cap W_\epsilon^u(-v) \neq \varnothing$, $W_\epsilon^s(-v) \cap W_\epsilon^u(v) \neq \varnothing$, and both intersections have two-sided crossings for ϵ sufficiently small. Then by Theorem 8.1, $\{P_\epsilon\}$ has positive topological entropy for ϵ sufficiently small, and by Abramov's Theorem [1], the geodesic flow for $\{g_\epsilon\}$ has positive topological entropy for ϵ sufficiently small.

Acknowledgment. The second author would like to thank the SFB 170 Göttingen for their generous support and hospitality during the summer of 1990. The authors would like to thank Keith Burns for many helpful discussions and comments. They would also like to thank Tasso Kaper for helpful conversations on the Poincaré/Melnikov Method.

Appendix

In this appendix we provide a short proof that for an ellipsoid with three different axes the geodesics passing through the umbilic points are asymptotic to the middle length basic closed geodesic. This is a folklore theorem whose statement appears several times in the literature, for example [15], [23]. Unfortunately, we have been unsuccessful in locating a proof, so for the reader's convenience, we present the following argument.

We use the notation from §2. We denote the shortest, middle length, and longest basic closed geodesic (as curves on E) by s, c, and l respectively. As in §2, we use the coordinate system $(u_1, u_2) \in (a_0, a_1) \times (a_1, a_2)$. We recall that this coordinate system is one to one in each of the eight regions bounded by the coordinate planes.

Lemma 1. *Suppose that $\sigma : [a, b] \to E$ is a geodesic from the invariant torus $F = \gamma$, $\gamma \in (a_0, a_1) \cup (a_1, a_2)$, and $(u_1(t), u_2(t))$ are the coordinates of $\sigma(t)$. Then*

$$\int_a^b \left| \frac{u_1(t)}{\sqrt{P_\gamma(u_1(t))}} \dot{u}_1(t) \right| dt = \int_a^b \left| \frac{u_2(t)}{\sqrt{P_\gamma(u_2(t))}} \dot{u}_2(t) \right| dt,$$

where $P_\gamma(x) = -x(x - a_0)(x - a_1)(x - a_2)(x - \gamma)$.

Proof. As stated in §2 the geodesics belonging to the invariant torus $F = \gamma$ are subject to the relation

$$\frac{\sqrt{U_1}}{\sqrt{-u_1 + \gamma}} \dot{u}_1 \pm \frac{\sqrt{U_2}}{\sqrt{u_2 - \gamma}} \dot{u}_2 = 0,$$

where $U_i = U_i(u_i) = (-1)^i u_i / f(u_i)$; $f(u_i) = 4(a_0 - u_i)(a_1 - u_i)(a_2 - u_i)$. Integrating this equation yields the proof of Lemma 1.

Lemma 2. *Consider the parameter $\gamma \in (a_1, a_2)$. Then the following equality holds*:

$$\int_\gamma^{a_2} \frac{x}{\sqrt{P_\gamma(x)}} dx = \int_{a_0}^{a_1} \frac{x}{\sqrt{P_\gamma(x)}} dx - \int_{-\infty}^{0} \frac{x}{\sqrt{P_\gamma(x)}} dx.$$

Proof. Consider the Riemann surface associated to the algebraic equation $w^2 = P_\gamma(z)$. This is a surface of genus 2 on which $\omega = (z/\sqrt{P_\gamma(z)})\, dz$ is a holomorphic differential. Integrating the differential along a suitable null homotopic path, we obtain the relation in Lemma 2.

Corollary 3. *Suppose that σ_γ is a geodesic on the invariant torus $F = \gamma \in (a_1, a_2)$. Let a and b be consecutive times at which σ_γ crosses l. Then the u_2-coordinate of $\sigma_\gamma(t)$ takes on each of the values γ and a_2 at most once for $t \in [a, b]$.*

Proof. It is well known that if $\gamma \in (a_1, a_2)$, then the geodesics on the torus $F = \gamma$ are always transverse to the lines of curvature $u_1 = $ const, and oscillate between the lines of curvature $u_2 = $ const, becoming tangent to them only when $u_2 = \gamma$. We have (cf. the proof of Proposition 4)

$$\int_a^b \left| \frac{u_1(t)}{\sqrt{P_\gamma(u_1(t))}} \dot{u}_1(t) \right| dt = 2\int_{a_0}^{a_1} \frac{x}{\sqrt{P_\gamma(x)}}\, dx.$$

If the lemma were false, we would have

$$\int_a^b \left| \frac{u_2(t)}{\sqrt{P_\gamma(u_2(t))}} \dot{u}_2(t) \right| dt \geq 2\int_{\gamma}^{a_2} \frac{x}{\sqrt{P_\gamma(x)}}\, dx.$$

Lemmas 1 and 2 show that this is impossible. q.e.d.

We now define a return map $R\colon [a_1, a_2] \to [a_1, a_2]$ that describes the behaviour of the geodesics passing through the umbilics. Choose an umbilic q. For $u_2 \in [a_1, a_2]$, let p denote one of the points on the longest basic closed geodesic l with coordinates (a_0, u_2) that is on the opposite side of the shortest closed geodesic s from q (or on s if $u_2 = a_2$). Define $R(u_2)$ so that $(a_0, R(u_2))$ are the coordinates of the point where the geodesic σ from p to q next intersects l. By the symmetry of the ellipsoid, this definition does not depend upon the choices that have been made. Obviously a_1 is a fixed point of R. The following proposition shows that it is the only fixed point.

Proposition 4. *For $u_2 \in (a_1, a_2]$ we have $R(u_2) < u_2$.*

Proof. Fix $u_2 \in (a_1, a_2]$ and let p, q and σ be as above, and parametrize σ so that $\sigma(0) = p$. For $\gamma \in (a_1, a_2)$ consider a family of geodesics σ_γ such that $\lim_{\gamma \to a_1} \sigma_\gamma = \sigma$ and each σ_γ is a geodesic on the invariant torus $F = \gamma$ with $\sigma_\gamma(0) = \sigma(0)$. It is easily seen that if γ is close to a_1, then σ_γ becomes tangent to the line of curvature $u_2 = \gamma$ at a point q_γ near q. Let t_γ be the first time after 0 when σ_γ crosses l, and let $v(\gamma)$ be the u_2-coordinate of $\sigma_\gamma(t_\gamma)$. During $[0, t_\gamma]$, the first coordinate $u_1(t)$ of σ_γ increases form a_0 to a maximum of a_1 when γ crosses c and then decreases to a_0. See Figure 1. Thus

$$\int_0^{t_\gamma} \left| \frac{u_1(t)}{\sqrt{P_\gamma(u_1(t))}} \dot{u}_1(t) \right| dt = 2\int_{\gamma}^{a_2} \frac{x}{\sqrt{P_\gamma(x)}}\, dx.$$

The second coordinate $u_2(t)$ satisfies $\gamma \leq u_2(t) \leq a_2$ and has the following values: u_2 at p; a_2 when σ_γ crosses s; γ at q_γ; and $v(\gamma)$ at time t_γ. We see from Corollary 3 that

$$\begin{aligned}\int_a^b \left| \frac{u_2(t)}{\sqrt{P_\gamma(u_2(t))}} \dot{u}_2(t) \right| dt &= \int_{u_2}^{a_2} \frac{x}{\sqrt{P_\gamma(x)}} dx + \int_\gamma^{a_2} \frac{x}{\sqrt{P_\gamma(x)}} dx \\ &\qquad + \int_\gamma^{v(\gamma)} \frac{x}{\sqrt{P_\gamma(x)}} dx \\ &= 2 \int_\gamma^{a_2} \frac{x}{\sqrt{P_\gamma(x)}} dx - \int_{v(\gamma)}^{u_2} \frac{x}{\sqrt{P_\gamma(x)}} dx.\end{aligned}$$

It follows from Lemmas 1 and 2 that

$$\int_{v(\gamma)}^{u_2} \frac{x}{\sqrt{P_\gamma(x)}} dx = \int_{-\infty}^{0} \frac{-x}{\sqrt{P_\gamma(x)}} dx.$$

Since the integrands on both sides are positive, this is possible only if $v(\gamma) < u_2$. Furthermore, if γ decreases, the integrand on the left decreases and the integrand on the right increases, which forces $v(\gamma)$ to decrease. Since $R(u_2) = \lim_{\gamma \to a_1} v(\gamma)$, we obtain the desired inequality.

Theorem 5. *Every geodesic that passes through an umbilic point is asymptotic to c.*

Proof. Every geodesic which passes through an umbilic point also passes through a pair of diametrical umbilics at fixed intervals. With at most one possible exception, every time it passes through one of the two regions bounded by l, it intersects s, and it follows by the proposition that it must be asymptotic to c.

References

[1] L. Abramov, *On the entropy of a flow,* Amer. Math. Soc. Transl. (2) **49** (1966) 167–170.

[2] H. Alkier, *Über geodätische Linien auf Fläschen zweiten Grades,* Dissertation Leipzig.

[3] K. Burns & M. Gerber, *Real analytic Bernoulli geodesic flows on S^2*, Ergodic Theory Dynamical Systems **9** (1989) 27–45 .

[4] C. Conley, *Twist mappings, linking, analyticity and periodic solutions which pass close to an unstable periodic solution,* Topological Dynamics (Joseph Auslander, ed.), Benjamin, New York, 1968.

[5] I. Cornfeld, S. Fomin &Y. Sinai, *Ergodic theory,* Springer, Berlin, 1982.

[6] W. de Melo, *Lectures on one-dimensional dynamics,* 17th Colóq. Brasileiro Mat., Inst. Mat. Pura Apl., 1990.

[7] V. Donnay, *Geodesic flow on the two-sphere,* I: *Positive measure entropy,* Ergodic Theory Dynamical Systems **8** (1989) 531–553 .

[8] ——, *Perturbations of elliptic billiards*, preprint.

[9] M Hirsch, C. Pugh & M. Shub, *Invariant manifolds*, Lecture Notes in Math., Vol. 583, Springer, Berlin, 1975.

[10] C. G. J. Jacobi, *Note von der geodätischen Linie auf einem Ellipsoid and der verschiedenen Anwendungen einer merkwürdigen analytischen Substitution*, Crelles J. **19** (1839) 309–313.

[11] A. Katok, *Ergodic perturbations of degenerate integrable Hamiltonian systems*, Math. USSR-Izv. 7 (3) (1973) 535–571 .

[12] ——, *Bernoulli diffeomorphisms on surfaces*, Ann. of Math. (2) **110** (1977) 529–574 .

[13] ——, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Études Sci. Inst. Publ. Hautes Math. **51** (1980) 137-173 .

[14] ——, *Nonuniform hyperbolicity and structure of smooth dynamical systems*, Proc. Internat. Congress Math. (Warszawa), 1983, 1245–1254 .

[15] W. Klingenberg, *Riemannian geometry*, de Gruyter, Berlin, 1982.

[16] S. Newhouse, *Continuity properties of entropy*, Ergodic Theory Dynamical Systems **8** (1988) 283–300 .

[17] H. Poincaré, *Sur les équations de la dynamique et le problème de trois corps*, Acta Math. **13** (1890) 1–270 .

[18] ——, *Les méthodes nouvelles de la mécanique céleste*, Dover, New York, 1957.

[19] C. Robinson, *Bifurcation to infinitely many sinks*, Comm. Math. Phys. **90** (1983) 433–459.

[20] H. Scüth, *Stabilität von periodischen Geodätischen auf n-dimensionalen Ellipsoiden*, Bonner Math. Schriften **60** (1972).

[21] S. Smale, *Diffeomorphisms with many periodic points*, Differential and Combinatorical Topology (S.S. Cairnes, ed.), Princeton University Press, Princeton, NJ, 1965, pp. 63-80.

[22] A. Thimm, *Integrabilität beim geodätischen Fluss*, Bonner Math. Schriften **103** (1978).

[23] H. Viesel, *Über einfach geschlossen Geodäische auf dem Ellipsoid*, Arch. Math. **22** (1971) 106–112.

[24] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Math., Vol. 79, Springer, Berlin, 1982.

[25] H. Weiss, *Surface diffeomorphisms having homoclinic points with 2-sided crossings have positive topological entropy*, preprint.

[26] Y. Yomdin, *Volume growth and entropy*, Israel J. Math. **57** (1987) 287–300.

UNIVERSITÄT AUGSBURG
PENNSYLVANIA STATE UNIVERSITY