ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

WEISHENG WU

On the ergodicity of geodesic flows on surfaces of nonpositive curvature

Tome XXIV, nº 3 (2015), p. 625-639.

<http://afst.cedram.org/item?id=AFST_2015_6_24_3_625_0>

Annales de la Faculté des Sciences de Toulouse Vol. XXIV, n° 3, 2015
pp. 625-639

On the ergodicity of geodesic flows on surfaces of nonpositive curvature

Weisheng Wu[1]

Résumé. — Soit M une surface lisse compacte de courbure négative ou nulle, de genre $\geqslant 2$. Nous prouvons l'ergodicité du flot géodésique sur le fibré tangent unitaire de M par rapport à la mesure de Liouville, en supposant que l'ensemble des points où la courbure est strictement négative a un nombre fini de composantes connexes. Sous la même hypothèse, nous prouvons qu'il n'existe pas de géodésique « plate » non-fermée. De plus, il existe au plus un nombre fini de bandes plates, et au plus un nombre fini de géodésiques fermées « plates » isolées.

Abstract. — Let M be a smooth compact surface of nonpositive curvature, with genus $\geqslant 2$. We prove the ergodicity of the geodesic flow on the unit tangent bundle of M with respect to the Liouville measure under the condition that the set of points with negative curvature on M has finitely many connected components. Under the same condition, we prove that a non-closed "flat" geodesic doesn't exist, and moreover, there are at most finitely many flat strips, and at most finitely many isolated closed "flat" geodesics.

1. Introduction

Let M be a smooth, connected and compact surface without boundary, with genus $g \geqslant 2$, and of nonpositive curvature. The geodesic flow Φ^t is defined on the unit tangent bundle T^1M. It is well known that the geodesic flow is Anosov when the curvature of the surface is strictly negative, and its ergodicity with respect to the Liouville measure ν can be proved by the

(*) Reçu le 10/10/2013, accepté le 05/05/2015

(1) School of Mathematical Sciences, Peking University, Beijing, 100871, China
wuweisheng@math.pku.edu.cn

Article proposé par Jean-Pierre Otal.

Hopf argument (cf., for example [2]). However, for surfaces of nonpositive curvature, the ergodicity of the geodesic flow is not known yet. The dynamical behavior of the flow gets more complicated because of the existence of the "flat geodesics" defined as follows. We define:

$$\Lambda := \{x \in T^1M : K(\gamma_x(t)) \equiv 0,\ \forall t \in \mathbb{R}\}$$

where K denotes the curvature of the point, and $\gamma_x(t)$ denotes the unique geodesic on M with an initial velocity $x \in T^1M$. We call γ_x a *flat geodesic* if $x \in \Lambda$, i.e., the curvature along the geodesic is always zero. It is proved that the geodesic flow is Anosov if and only if $\Lambda = \emptyset$ (cf. [6]), and in this case the ergodicity follows from the Hopf argument.

By Pesin's well-known result (cf. [1]), the geodesic flow is ergodic on the following set:

$$\Delta := \{x \in T^1M : \limsup_{t\to\infty} \frac{1}{t}\int_0^t K(\gamma_x(s))ds < 0\}. \tag{1.1}$$

Clearly $\Delta \subset \Lambda^c$. It is stated in [4] that the geodesic flow is also ergodic on Λ^c. Indeed, we have

Lemma 1.1. — $\nu(\Lambda^c \setminus \Delta) = 0$.

Proof. — Assume $\nu(\Lambda^c \setminus \Delta) > 0$. Let $\pi : T^1M \to M$ be the natural projection. Denote $f(x) := \chi_{\Lambda^c\setminus\Delta}(x)\cdot K(\pi(x))$. Note that $f(x) \leqslant 0$. By Birkhoff ergodic theorem, for ν-a.e. $x \in T^1M$,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\Phi^s(x))ds := \tilde{f}(x)$$

and

$$\int_{T^1M} \tilde{f}(x)d\nu(x) = \int_{T^1M} f(x)d\nu(x) \leqslant 0. \tag{1.2}$$

By the definition of Δ in (1.1), $\tilde{f}(x) = 0$ for ν-a.e. $x \in T^1M$. Then by (1.2), $\int_{T^1M} f(x)d\nu(x) = 0$, so $f(x) = 0$ for ν-a.e. $x \in T^1M$. Hence, $K(\pi(x)) = 0$ for ν-a.e. $x \in \Lambda^c \setminus \Delta$. Since the orbit foliation of Φ^t is smooth, for ν-a.e. $x \in \Lambda^c \setminus \Delta$, one has $K(\Phi^t(x)) = 0$ for a.e. t. By continuity of the curvature function K, we have $K(\Phi^t(x)) \equiv 0$ for $\forall t \in \mathbb{R}$, i.e., $x \in \Lambda$, a contradiction to $x \in \Lambda^c \setminus \Delta$. Therefore, $\nu(\Lambda^c \setminus \Delta) = 0$. □

So the geodesic flow is ergodic on the set Λ^c. Therefore, the geodesic flow is ergodic on T^1M if $\nu(\Lambda) = 0$. It is not known in general if $\nu(\Lambda) = 0$, but this is the case for all the known examples so far. Moreover, in all these examples, the flat geodesics are always closed. This motivates the following conjecture whose statement is stronger than ergodicity:

CONJECTURE 1.2. *(cf.[10]). — All flat geodesics are closed and there are only finitely many homotopy classes of such geodesics. In particular,* $\nu(\Lambda) = 0$ *and hence the geodesic flow is ergodic.*

In this paper we prove the following two theorems according to the dichotomy: (1) $\Lambda \subset \mathrm{Per}(\Phi)$; (2)$\Lambda \cap (\mathrm{Per}\ (\Phi))^c \neq \emptyset$. Here $\mathrm{Per}(\Phi)$ denotes the set of periodic points of the geodesic flow, and $\mathcal{O}(z)$ denotes the orbit of z under the geodesic flow.

THEOREM 1.3. — *If* $\Lambda \subset Per(\Phi)$, *then*

$$\Lambda = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \ldots \mathcal{O}_k \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \ldots \cup \mathcal{F}_l,$$

where each $\mathcal{O}_i, 1 \leqslant i \leqslant k$ *is an isolated periodic orbit and each* $\mathcal{F}_j, 1 \leqslant j \leqslant l$ *consists of vectors tangent to a flat strip. Here* k *or* l *are allowed to be* 0 *if there is no isolated closed flat geodesic or no flat strip.*

THEOREM 1.4. — *If* $\Lambda \cap (Per\ (\Phi))^c \neq \emptyset$, *then there exist* $y, z \in \Lambda$, $y \notin \mathcal{O}(z)$, *such that*

$$d(\Phi^t(y), \Phi^t(z)) \to 0, \quad \text{as } t \to +\infty.$$

In the process of proving the above two theorems, we also obtain a result of independent importance:

THEOREM 1.5. — $\Lambda \cap (Per\ (\Phi))^c$ *is a closed set in* Λ.

Theorem 1.5 says that if we count vectors tangent to a flat strip as a single orbit, then closed flat orbits must be isolated from non-closed flat orbits.

Let $\{p \in M : K(p) < 0\}$ be the set of points with negative curvature on M. As a consequence of Theorem 1.3 and 1.4, we can prove Conjecture 1.2 in the case when $\{p \in M : K(p) < 0\}$ has only finitely many connected components:

THEOREM 1.6. — *If the set* $\{p \in M : K(p) < 0\}$ *has finitely many connected components, then* $\Lambda \subset Per(\Phi)$. *In particular, the geodesic flow is ergodic.*

Theorem 1.6 gives a negative answer to Question 6.2.1 asked by Burns in a recent survey [4], for the case when $\{p \in M : K(p) < 0\}$ has only finitely many connected components. Furthermore, by Theorem 1.3 there are at most finitely many flat strips and isolated closed flat geodesics in this case. But we don't know the answer to Question 6.2.1 in [4] for the general case.

The paper is organized as follows. In section 2, we present some preliminaries and well known results. The proofs of Theorems 1.3, 1.4 and 1.5 will occupy Section 3. In Section 4, we prove Theorem 1.6 and ask a further related question.

2. Preliminaries

2.1. Universal Cover

Consider the universal covering space $\tilde{M}$ of M, which can be identified with the unit disk in the plane. The lifting of a geodesic γ from M to $\tilde{M}$ is denoted by $\tilde{\gamma}$. All the geodesics are supposed to have unit speed. It is well known that $\tilde{M}$ is a Hadamard manifold with many nice properties. For any two given points in $\tilde{M}$, there exists a unique geodesic joining them. Two geodesics $\tilde{\gamma_1}$ and $\tilde{\gamma_2}$ are said to be asymptotes if $d(\tilde{\gamma_1}(t), \tilde{\gamma_2}(t)) \leqslant C$ for some $C > 0$ and $\forall t > 0$. This relation is an equivalence relation. Denote by $\tilde{M}(\infty)$ the set of all equivalence classes, which can be identified with the boundary of the unit disk. We denote by $\tilde{\gamma}(+\infty)$ the class of the geodesic $\tilde{\gamma}$, and by $\tilde{\gamma}(-\infty)$ the one of the reversed geodesic to $\tilde{\gamma}$.

Any closed geodesic γ in M can be lifted to a geodesic $\tilde{\gamma}$ on $\tilde{M}$, such that

$$\tilde{\gamma}(t + t_0) = \phi(\tilde{\gamma}(t)), \quad \forall t \in \mathbb{R}$$

for some $t_0 > 0$ and $\phi \in \pi_1(M)$. In this case, we say that ϕ fixes $\tilde{\gamma}$, i.e., $\phi(\tilde{\gamma}) = \tilde{\gamma}$. Then ϕ acts on $\tilde{M}(\infty)$ in the natural way and fixes exactly two points $\tilde{\gamma}(\pm\infty)$. Moreover for any $x \in \tilde{M}(\infty)$ and $x \neq \tilde{\gamma}(\pm\infty)$, we have $\lim_{n\to+\infty} \phi^n(x) = \tilde{\gamma}(+\infty)$ and $\lim_{n\to-\infty} \phi^n(x) = \tilde{\gamma}(-\infty)$.

There are two continuous one dimensional distributions E^s and E^u on T^1M which are invariant under the derivative of Φ^t (cf. [7]). Their integral manifolds form foliations W^s and W^u of T^1M respectively which are invariant under Φ^t, known as the stable and unstable horocycle foliations. The lifting of W^s and W^u to $T^1\tilde{M}$ are denoted by $\tilde{W}^s$ and $\tilde{W}^u$ respectively. If $w \in \tilde{W}^s(v)$, then geodesics $\tilde{\gamma}_v(t)$ and $\tilde{\gamma}_w(t)$ are asymptotic.

2.2. Area of ideal triangles

Given $x, y, z \in \tilde{M}(\infty)$, an ideal triangle with vertices x, y, z means the region in $\tilde{M}$ bounded by the three geodesics joining x and y, y and z, z and

x. It is an interesting topic to study the area of ideal triangles. We have the following theorem due to R. Ruggiero:

THEOREM 2.1. (cf. [11]). — *If $K(\tilde{\gamma}(t)) \equiv 0$ for $\forall t \in \mathbb{R}$, then every ideal triangle having $\tilde{\gamma}(t)$ as an edge has infinite area.*

In fact, if we have a triangle with vertices x, a, b, where $x = \tilde{\gamma}_1(+\infty) = \tilde{\gamma}_2(+\infty)$, $a \in \tilde{\gamma}_1$, $b \in \tilde{\gamma}_2$, and $\tilde{\gamma}_1$ is a flat geodesic, then the triangle has infinite area. The proof follows from the fact that the length of a stable Jacobi fields decreases slowly along a geodesic with curvature close to zero (cf. [11]).

2.3. Flat strips

A flat strip means a totally geodesic isometric imbedding $r : \mathbb{R} \times [0, c] \to \tilde{M}$, where $\mathbb{R} \times [0, c]$ is a strip in an Euclidean plane. We have the following flat strip lemma due to P. Eberlein and B. O'Neill:

LEMMA 2.2. (cf. [8]). — *If two distinct geodesics $\tilde{\alpha}$ and $\tilde{\beta}$ satisfy $d(\tilde{\alpha}(t), \tilde{\beta}(t)) < C$ for some $C > 0$ and $\forall t \in \mathbb{R}$, then they are the boundary curves of a flat strip in $\tilde{M}$.*

We also call the projection of a flat strip onto M a flat strip. An important progress toward Conjecture 1.2 was made by J. Cao and F. Xavier on the flat geodesics inside flat strips:

THEOREM 2.3. (cf. [5]). — *A flat strip on M consists of closed flat geodesics in the same homotopy type.*

3. Main Construction

In this section, we mainly carry out two constructions based on a similar idea. First, we prove Theorem 1.4 by constructing two points y, z with the required property in the theorem starting from an aperiodic orbit of $x \in \Lambda$. Second, assume the contrary for Theorem 1.3, i.e., there exist infinitely many periodic orbits, then we can construct an aperiodic orbit starting from them. Both constructions are based on the *expansivity* property:

DEFINITION 3.1. (cf. Definition 3.2.11 in [9]). — *$x \in T^1M$ has the expansivity property if there exists a small $\delta_0 > 0$, such that if $d(\Phi^t(x), \Phi^t(y) < \delta_0$ for $\forall t \in \mathbb{R}$, then $y = \Phi^{t_0}(x)$ for some t_0 with $|t_0| < \delta_0$.*

LEMMA 3.2. — *If x is not tangent to a flat strip, it has the expansivity property.*

Proof. — Assume not. Then for an arbitrarily small $\epsilon > 0$ less than the injectivity radius of M, there exists y such that $y \notin \mathcal{O}(x)$ and $d(\gamma_x(t), \gamma_y(t)) < \epsilon$ for $\forall t \in \mathbb{R}$. By the choice of ϵ, we can lift $\gamma_x(t)$ and $\gamma_y(t)$ to the universal cover $\tilde{M}$ such that

$$d(\tilde{\gamma}_x(t), \tilde{\gamma}_y(t)) < \epsilon \quad \text{for } \forall t \in \mathbb{R}.$$

Thus by Lemma 2.2, $\tilde{\gamma}_x(t)$ and $\tilde{\gamma}_y(t)$ bound a flat strip. Hence x is tangent to a flat strip, a contradiction. $\square$

We prove Theorem 1.4 and Theorem 1.5 in the next subsection. After that we prove Theorem 1.3.

3.1. Proof of Theorem 1.4 and 1.5

Now we assume that $\Lambda \cap (\text{Per } \Phi)^c \neq \emptyset$, in other words, there exists an aperiodic orbit $\mathcal{O}(x)$ in Λ. We will construct y, z as in Theorem 1.4 starting from $\mathcal{O}(x)$. First we can always find two arbitrarily close points on the orbit $\mathcal{O}(x)$:

Lemma 3.3. — *For any $k \in \mathbb{N}$, there exist two sequences $t_k \to +\infty$, and $t'_k \to +\infty$ such that $t'_k - t_k \to +\infty$ and*

$$d(x_k, x'_k) < \frac{1}{k}, \qquad \textit{where} \ \ x_k = \Phi^{t_k}(x), \ x'_k = \Phi^{t'_k}(x).$$

Proof. — For any fixed $k \in \mathbb{N}$, let $\epsilon < \frac{1}{2k}$ be sufficiently small. We choose a segment $[z_k, w_k]$ along the orbit $\mathcal{O}(x)$ from z_k to w_k with length T_k. Let X be the vector field tangent to the geodesic flow on T^1M, and $X^\perp$ be the orthogonal complement of X, i.e. a two dimensional smooth distribution on T^1M. For any $y \in [z_k, w_k]$ define $D_\epsilon(y) := \exp_y(X^\perp_\epsilon(y))$, where $X^\perp_\epsilon(y)$ denotes the ϵ-ball centered at origin in the subspace $X^\perp(y)$.

Assume $D_\epsilon(z) \cap D_\epsilon(w) = \emptyset$ for any $z, w \in [z_k, w_k]$. Since T^1M is compact and its curvature is bounded, we have the following estimates on the volume:

$$C_0 \epsilon^2 T_k \leqslant \text{Vol}(\bigcup_{y \in [z_k, w_k]} D_\epsilon(y)) \leqslant \text{Vol}(T^1M).$$

But the above inequalities don't hold if we choose T_k large enough. So there are two points in $[z_k, w_k]$, say, x_k, x'_k such that $D_\epsilon(x_k) \cap D_\epsilon(x'_k) \neq \emptyset$, and hence $d(x_k, x'_k) < 2\epsilon < \frac{1}{k}$. Let $x_k = \Phi^{t_k}(x)$, $x'_k = \Phi^{t'_k}(x)$ where we can make $t'_k - t_k \to +\infty$ as $k \to +\infty$. $\square$

For any pair of x_k, x'_k with large enough k, we claim the expansivity in the positive direction of the flow:

Proposition 3.4. — *Fix an arbitrarily small $\epsilon_0 > 0$. There exists $s_k \to +\infty$, such that*

$$d(\Phi^{s_k}(x_k), \Phi^{s_k}(x'_k)) = \epsilon_0,$$

$$\text{and } \ d(\Phi^s(x_k), \Phi^s(x'_k)) < \epsilon_0 \quad \text{for } \forall\ 0 \leqslant s < s_k.$$

Remark 3.5. — In fact for the purpose of our construction, it is enough to have the expansivity in either positive or negative direction. And this is easily known since x is not periodic and hence not tangent to a flat strip by Theorem 2.3. But in Proposition 3.4, we have a stronger statement that the flow is expansive in the positive direction. To prove it, we will make use of several lemmas which seem to be of independent interest.

The following lemma was proved in [3] and stated in [5]:

Lemma 3.6. (cf. [3], [5]). — *If $w' \in W^s(w)$ and $\lim_{t\to+\infty} d(\gamma_w(t), \gamma_{w'}(t)) = \delta > 0$, then $\gamma_w(t)$ and $\gamma_{w'}(t)$ converge to the boundaries of a flat strip of width δ.*

Proof. — Suppose that $\lim_{s_i\to+\infty} \Phi^{s_i}(w) = v$ and $\lim_{s_i\to+\infty} \Phi^{s_i}(w') = v'$, then $v' \in W^s(v)$ and for any $t \in \mathbb{R}$:

$$d(\gamma_v(t), \gamma_{v'}(t)) = \lim_{s_i\to+\infty} d(\gamma_w(t+s_i), \gamma_{w'}(t+s_i)) = \delta.$$

Hence we can lift the geodesics to $\tilde{M}$ such that $v' \in \tilde{W}^s(v)$ and

$$d(\tilde{\gamma}_v(t), \tilde{\gamma}_{v'}(t)) = \delta \quad \text{for } \forall t \in \mathbb{R}.$$

By Lemma 2.2, $\tilde{\gamma}_v(t)$ and $\tilde{\gamma}_{v'}(t)$ are the boundaries of a flat strip of width δ. □

The next lemma says that a flat geodesic converges to a closed geodesic (no matter flat or not), then the former must be closed as well and coincide with the latter.

Lemma 3.7. — *Suppose that $y \in \Lambda$, and the ω-limit set $\omega(y) = \mathcal{O}(z)$ where $\mathcal{O}(z)$ is periodic. Then $\mathcal{O}(y) = \mathcal{O}(z)$. In particular, $\mathcal{O}(y)$ is periodic.*

Proof. — First we prove that we can lift geodesics $\gamma_z(t), \gamma_y(t)$ to the universal cover $\tilde{M}$, denoted by $\tilde{\gamma}_0(t)$ and $\tilde{\gamma}(t)$ respectively, such that $\tilde{\gamma}_0(+\infty) = \tilde{\gamma}(+\infty)$. But this is guaranteed by the assumption $\omega(y) = \mathcal{O}(z)$. Moreover, $\lim_{t\to+\infty} d(\tilde{\gamma}_0(t), \tilde{\gamma}(t)) = 0$.

Since $\gamma_z(t)$ is a closed geodesic, there exists an isometry ϕ of $\tilde{M}$ such that $\phi(\tilde{\gamma}_0(t)) = \tilde{\gamma}_0(t+t_0)$. Moreover, on the boundary of the disk $\tilde{M}(\infty)$,

ϕ fixes exactly two points $\tilde{\gamma}_0(\pm\infty)$, and for any other point $a \in \tilde{M}(\infty)$, $\lim_{n\to+\infty} \phi^n(a) = \tilde{\gamma}_0(+\infty)$.

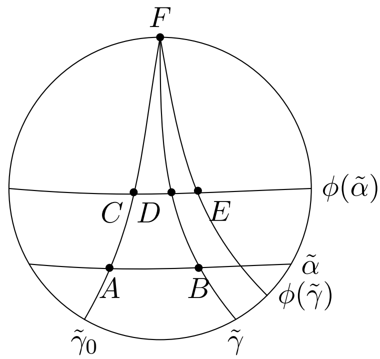

Figure 1. — Proof of Lemma 3.7

Assume $\tilde{\gamma}$ is not fixed by ϕ. Then $\tilde{\gamma}$ and $\phi(\tilde{\gamma})$ don't intersect since $\phi(\tilde{\gamma})(+\infty) = \tilde{\gamma}(+\infty)$. We pick another geodesic $\tilde{\alpha}$ as shown in Figure 1. The image of infinite triangle ABF under ϕ is the infinite triangle CEF. Since ϕ is an isometry, it preserves area. With a limit process, it is easy to show that Area of $ABCD \geqslant$ Area of DEF. But since $\tilde{\gamma}$ is a flat geodesic, Area of DEF is infinite by Theorem 2.1, which is a contradiction to the fact that $ABCD$ has finite area. So $\phi(\tilde{\gamma})$ and $\tilde{\gamma}$ must coincide.

Hence $\tilde{\gamma}(\pm\infty) = \tilde{\gamma}_0(\pm\infty)$. Then either $\tilde{\gamma}(t)$ and $\tilde{\gamma}_0(t)$ bound a flat strip by Lemma 2.2 or $\tilde{\gamma}(t) = \tilde{\gamma}_0(t)$. Since $\lim_{t\to+\infty} d(\tilde{\gamma}(t), \tilde{\gamma}_0(t)) = 0$, we have $\tilde{\gamma}(t) = \tilde{\gamma}_0(t)$. Hence $\mathcal{O}(y) = \mathcal{O}(z)$. $\square$

We improve Lemma 3.7 as follows.

Lemma 3.8. — *Suppose that $y \in \Lambda$ and $z \in \omega(y)$ where z is periodic. Then $\mathcal{O}(y) = \mathcal{O}(z)$. In particular, y is periodic.*

Proof. — Suppose that there exist $s_k \to +\infty$ such that $\Phi^{s_k}(y) \to z$. If $\Phi^{s_k}(y) \in W^s(z)$ for some k then we must have $\omega(y) = \mathcal{O}(z)$. Then by Lemma 3.7, we have $\mathcal{O}(\Phi^{s_k}(y)) = \mathcal{O}(z)$. So we are done.

Suppose that $\Phi^{s_k}(y) \notin W^s(z)$ for any k. Note that if $y \neq z$ then y and z can not be tangent to a same flat strip. Therefore, for any small $\epsilon_0 > 0$ and any large k, there exists a l_k with $l_k \to +\infty$ such that

$$d(\Phi^{l_k}(\Phi^{s_k}(y)), \Phi^{l_k}(z)) = \epsilon_0,$$

where we take l_k to be the smallest positive number to satisfy the above equality. By taking a subsequence but still using the same notation for simplicity, we assume that

$$\Phi^{l_k}(\Phi^{s_k}(y)) \to y^+, \quad \text{and} \quad \Phi^{l_k}(z) \to z^+ \tag{3.1}$$

as $k \to +\infty$. Then z^+ is periodic and $d(y^+, z^+) = \epsilon_0$. For any $t > 0$, since $0 < -t + l_k < l_k$ for large enough k, one has

$$d(\Phi^{-t}(y^+), \Phi^{-t}(z^+)) = \lim_{k\to+\infty} d(\Phi^{-t+l_k+s_k}(y)), \Phi^{-t+l_k}(z)) \leqslant \epsilon_0.$$

So $-y^+ \in W^s(-z^+)$. Replacing y, z by $-y, -z$ respectively and applying the same argument, we can obtain two points y^-, z^- such that $-y^- \in W^s(-z^-)$ and $d(y^-, z^-) = \epsilon_0$, $y^- \in \omega(-y)$ and z^- is periodic. Then we have the following three cases:

1. $\lim_{t\to\infty} d(\Phi^t(-y^+), \Phi^t(-z^+)) = 0$. By Lemma 3.7, $-y^+$ is periodic and in fact $-y^+ = -z^+$ as $\lim_{t\to\infty} d(\Phi^t(-y^+), \Phi^t(-z^+)) = 0$. This contradicts to $d(y^+, z^+) = \epsilon_0$.

2. $\lim_{t\to\infty} d(\Phi^t(-y^-), \Phi^t(-z^-)) = 0$. By Lemma 3.7, $-y^-$ is periodic and in fact $-y^- = -z^-$ as $\lim_{t\to\infty} d(\Phi^t(-y^-), \Phi^t(-z^-)) = 0$. This contradicts to $d(y^-, z^-) = \epsilon_0$.

3. $\lim_{t\to\infty} d(\Phi^t(-y^+), \Phi^t(-z^+)) = \delta_1$ and $\lim_{t\to\infty} d(\Phi^t(y^-), \Phi^t(z^-)) = \delta_2$ for some $\delta_1, \delta_2 > 0$. By Lemma 3.6 $-y^+$ converges to a closed flat geodesic. Then by Lemma 3.7 γ_{y^+} and γ_z are boundaries of a flat strip of width δ_1. By the same argument γ_{y^-} and γ_z are boundaries of a flat strip of width δ_2. We claim that these two flat strips lie on the different sides of γ_z. Indeed, we choose ϵ_0 small enough and consider the ϵ_0 neighborhood of the closed geodesic γ_z which contains two regions lying on the different sides of γ_z. We can choose the sequences in (3.1) for y and $-y$ respectively such that y^+ and y^- lie in different regions as above. This implies the claim. So we get a flat strip of width $\delta_1 + \delta_2$ and z is tangent to the interior of the flat strip. Now recall that $y^+ \in \omega(y)$ and y^+ is periodic, so we can apply all the arguments above to y^+ instead of z. Either we are arriving at a contradiction as in case (1) or case (2) and we are done, or we get a flat strip of width greater than $\delta_1 + \delta_2$. But we can not enlarge a flat strip again and again in a compact surface M. So we are done.

□

Proof of Theorem 1.5. — Assume that there exists a sequence $y_k \in \Lambda \cap (\mathrm{Per}\ (\Phi))^c$ such that $\lim_{k\to+\infty} y_k = z$ for some $z \in \Lambda \cap \mathrm{Per}\ (\Phi)$. We can apply the same argument in the proof of Lemma 3.8 replacing $\Phi^{s_k}(y)$ by y_k to get a contradiction. □

Proof of Proposition 3.4. — Assume the contrary, i.e. $d(\Phi^s(x_k), \Phi^s(x'_k)) \leqslant \epsilon_0$ for $\forall s > 0$. Then the two geodesics γ_{x_k} and $\gamma_{x'_k}$ are asymptotic. Without loss of generality, we suppose that $x'_k \in W^s(x_k)$. By the convexity of $d(\gamma_{x_k}(t), \gamma_{x'_k}(t))$, we have either $\lim_{t\to+\infty} d(\gamma_{x_k}(t), \gamma_{x'_k}(t)) = 0$ or $\lim_{t\to+\infty} d(\gamma_{x_k}(t), \gamma_{x'_k}(t)) = \delta > 0$.

- If $\lim_{t\to+\infty} d(\gamma_{x_k}(t), \gamma_{x'_k}(t)) = 0$, then we can choose a subsequence $s_i \to +\infty$, and z such that
$$\lim_{s_i\to+\infty} \Phi^{s_i}(x_k) = z$$
and
$$\lim_{s_i\to+\infty} \Phi^{s_i}(x'_k) = z.$$
Since $x_k = \Phi^{t_k}(x)$ and $x'_k = \Phi^{t'_k}(x)$ with $t'_k - t_k \to +\infty$ as $k \to \infty$, we have $\lim_{s_i\to+\infty} \Phi^{s_i}(x'_k) = \lim_{s_i\to+\infty} \Phi^{t'_k - t_k} \circ \Phi^{s_i}(x_k) = \Phi^{t'_k - t_k}(z)$. Hence $\Phi^{t'_k - t_k}(z) = z$, i.e., z is a periodic point in Λ. As $z \in \omega(x_k)$, x_k is periodic by Lemma 3.8. Hence x is periodic as well. But we assume x is aperiodic at the beginning. A contradiction.
- If $\lim_{t\to+\infty} d(\gamma_{x_k}(t), \gamma_{x'_k}(t)) = \delta > 0$, then $\omega(x_k) = \mathcal{O}(w)$ where w is tangent to a boundary of a flat strip by Lemma 3.6. Then w is periodic by Theorem 2.3. Hence x_k is periodic by Lemma 3.7. A contradiction.

In each case we arrive at a contradiction, so we are done. □

Now we continue with our construction.

Proposition 3.9. — *For arbitrarily small $\epsilon_0 > 0$, there exist $a, b \in \Lambda \cap (Per\ (\Phi))^c$ such that*
$$d(a, b) = \epsilon_0, \tag{3.2}$$
$$d(\Phi^t(a), \Phi^t(b)) \leqslant \epsilon_0 \qquad \forall t < 0, \tag{3.3}$$
$$a \notin \mathcal{O}(b), \tag{3.4}$$
$$a \in W^u(b). \tag{3.5}$$

Proof. — We apply Proposition 3.4. We can pick a subsequence $k_i \to +\infty$, such that
$$\lim_{k_i\to+\infty} \Phi^{s_{k_i}}(x_{k_i}) = a,$$
and
$$\lim_{k_i\to+\infty} \Phi^{s_{k_i}}(x'_{k_i}) = b.$$
Then $d(a, b) = \lim_{k_i\to+\infty} d(\Phi^{s_{k_i}}(x_{k_i}), \Phi^{s_{k_i}}(x'_{k_i})) = \epsilon_0$. We get (3.2).

For any $t < 0$, since $0 < s_{k_i} + t < s_{k_i}$ for large k_i, we have:

$$d(\Phi^t(a), \Phi^t(b)) = \lim_{k_i \to +\infty} (d(\Phi^{s_{k_i}+t}(x_{k_i}), \Phi^{s_{k_i}+t}(x'_{k_i}))) \leqslant \epsilon_0.$$

Hence we get (3.3).

Next assume that a is periodic. Since

$$\lim_{k_i \to +\infty} \Phi^{t_{k_i}+s_{k_i}}(x) = \lim_{k_i \to +\infty} \Phi^{s_{k_i}}(x_{k_i}) = a,$$

then x is periodic by Lemma 3.8. A contradiction. So $a \in (\mathrm{Per}\ (\Phi))^c$. Similarly $b \in (\mathrm{Per}\ (\Phi))^c$. Thus $a, b \in \Lambda \cap (\mathrm{Per}\ (\Phi))^c$.

Now we prove (3.4), i.e., $a \notin \mathcal{O}(b)$. For a simpler notation, we assume

$$\lim_{k \to +\infty} \Phi^{s_k}(x_k) = a,$$

and

$$\lim_{k \to +\infty} \Phi^{s_k}(x'_k) = b.$$

We can lift $\gamma_{x_k}(t), \gamma_{x'_k}(t)$ on M to geodesics $\tilde{\gamma}_k, \tilde{\gamma}'_k$ respectively on $\tilde{M}$ in the way such that $d(x_k, x'_k) < \frac{1}{k}$, $d(y_k, y'_k) = \epsilon_0$, where $y_k = \Phi^{s_k}(x_k)$, $y'_k = \Phi^{s_k}(x'_k)$, and moreover $y_k \to a$, $y'_k \to b$. Then $\tilde{\gamma}_k$ converges to $\tilde{\gamma} = \tilde{\gamma}_a$, $\tilde{\gamma}'_k$ converges to $\tilde{\gamma}' = \tilde{\gamma}_b$ and $d(a, b) = \epsilon_0$. See Figure 2 (we use same notation for a vector and its footpoint).

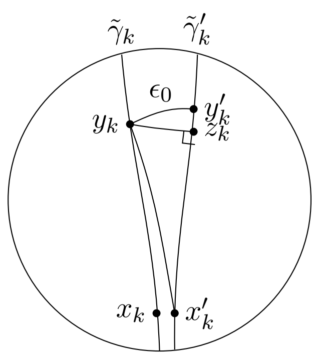

Figure 2. — Proof of $\tilde{\gamma} \neq \tilde{\gamma}'$

First we show that $d(y_k, \tilde{\gamma}'_k)$ is bounded away from 0. Write $d_k := d(y_k, \tilde{\gamma}'_k) = d(y_k, z_k)$, $l_k := d(y_k, x'_k)$, $b_k := d(x'_k, z_k)$, and $b'_k := d(z_k, y'_k)$. And we already know that $d(x'_k, y'_k) = s_k$. Suppose that $d_k \to 0$ as $k \to +\infty$. By triangle inequality, $\lim_{k \to +\infty}(l_k - b_k) = 0$. But since $\lim_{k \to +\infty}(l_k - s_k) \leqslant$

$\lim_{k\to+\infty} d(x_k, x'_k) = 0$, we have that $\lim_{k\to+\infty} b'_k = \lim_{k\to+\infty} |(l_k - b_k) - (l_k - s_k)| = 0$. But by triangle inequality, $\epsilon_0 \leqslant d_k + b'_k \to 0$. A contradiction. Now $\tilde{\gamma} \neq \tilde{\gamma}'$ follows from $d(a, \tilde{\gamma}') = \lim_{k\to+\infty} d(y_k, \tilde{\gamma}'_k) \geqslant d_0$ for some $d_0 > 0$.

Next we suppose there exists a $\phi \in \pi_1(M)$ such that $\phi(\tilde{\gamma}) = \tilde{\gamma}'$. See Figure 3. Observe that $\tilde{\gamma}(-\infty) = \tilde{\gamma}'(-\infty)$ since $d(\Phi^t(a), \Phi^t(b)) \leqslant \epsilon_0$, for $\forall t < 0$. Let $\tilde{\gamma}_0$ be the closed geodesic such that $\phi(\tilde{\gamma}_0) = \tilde{\gamma}_0$. Then $\tilde{\gamma}(-\infty) = \tilde{\gamma}_0(-\infty)$. By Lemma 3.7, $\tilde{\gamma}$ is a closed geodesic, i.e., a is a periodic point. We arrive at a contradiction. Hence for any $\phi \in \pi_1(M)$, $\phi(\tilde{\gamma}) \neq \tilde{\gamma}'$. So $a \notin \mathcal{O}(b)$, and we get (3.4).

At last, if $a \notin W^u(b)$, we can replace a by some $a' \in \mathcal{O}(a)$, b by some $b' \in \mathcal{O}(b)$ such that $a' \in W^u(b')$ and the above three properties still hold for a different ϵ_0. We get (3.5).

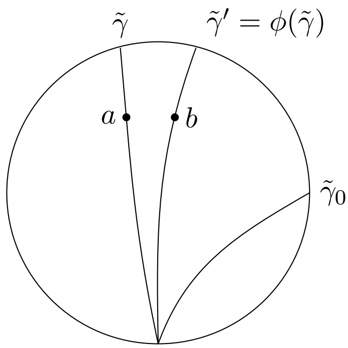

Figure 3. — Proof of $\phi(\tilde{\gamma}) \neq \tilde{\gamma}'$

Proof of Theorem 1.4. — We apply Proposition 3.9. Let $y = -a$, $z = -b$, then $y, z \in \Lambda \cap (\text{Per }(\Phi))^c$, $d(\Phi^t(y), \Phi^t(z)) \leqslant \epsilon_0$, $\forall t > 0$, $z \notin \mathcal{O}(y)$ and $y \in W^s(z)$.

If ϵ_0 is small enough, we can lift geodesics $\gamma_y(t)$ and $\gamma_z(t)$ to $\tilde{\gamma}_y(t)$ and $\tilde{\gamma}_z(t)$ respectively on $\tilde{M}$ such that $d(\tilde{\gamma}_y(t), \tilde{\gamma}_z(t)) \leqslant \epsilon_0$ for any $t > 0$ and $y \in \tilde{W}^s(z)$. Suppose $\lim_{t\to+\infty} d(\tilde{\gamma}_y(t), \tilde{\gamma}_z(t)) = \delta > 0$. Then by Lemma 3.6, $\tilde{\gamma}_y(t)$ and $\tilde{\gamma}_z(t)$ converge to the boundaries of a flat strip. Hence y and z are periodic by Lemma 3.7. A contradiction. So $\lim_{t\to+\infty} d(\tilde{\gamma}_y(t), \tilde{\gamma}_z(t)) = 0$. Hence $d(\Phi^t(y), \Phi^t(z)) \to 0$, as $t \to +\infty$. ☐

3.2. Proof of Theorem 1.3

Part of the proof of Theorem 1.3 is a verbatim repetition of the one of Proposition 3.9, so we omit it.

Proof of Theorem 1.3. — Suppose that $\Lambda \subset \text{Per}\ (\Phi)$. If $x \in \Lambda$, then x is tangent to an isolated closed flat geodesic or a flat strip.

Assume the contrary. Then there exists a sequence of different vectors $x'_k \in \Lambda$ such that $\lim_{k\to+\infty} x'_k = x$ for some $x \in \Lambda$. Here different x'_k are tangent to different isolated closed geodesics or to different flat strips, and x is tangent to an isolated flat closed geodesic or to a flat strip. For large enough k, we suppose that $d(x'_k, x) < \frac{1}{k}$. Fix any small $\epsilon_0 > 0$. It is impossible that $d(\Phi^t(x'_k), \Phi^t(x)) \leqslant \epsilon_0$ for $\forall t > 0$. Otherwise, $\tilde{\gamma}_{x'_k}(t), \tilde{\gamma}_x(t)$ are positively asymptotic closed geodesics so they must tangent to a common flat strip by Lemma 3.6 and Lemma 3.7. This is impossible since different x'_k are tangent to different isolated closed flat geodesics or to different flat strips. Hence there exists a sequence $s_k \to +\infty$ such that

$$d(\Phi^{s_k}(x'_k), \Phi^{s_k}(x)) = \epsilon_0,$$

and

$$d(\Phi^s(x'_k), \Phi^s(x)) \leqslant \epsilon_0 \quad \forall 0 \leqslant s < s_k.$$

Write $y_k := \Phi^{s_k}(x)$ and $y'_k := \Phi^{s_k}(x'_k)$. Without loss of generality, suppose that $y_k \to a$ and $y'_k \to b$. A similar proof as in Proposition 3.9 gives $d(a, b) = \epsilon$ and $d(\Phi^t(a), \Phi^t(b)) \leqslant \epsilon_0$ for $\forall t \leqslant 0$. If we lift the geodesics to $\tilde{M}$ (using the same notation as in the proof of Proposition 3.9), we can prove $\tilde{\gamma} \neq \tilde{\gamma}'$ similarly. But then we have two closed flat geodesics $\tilde{\gamma}$ and $\tilde{\gamma}'$ that are negatively asymptotic, so they must coincide by Lemma 3.7. A contradiction. □

4. Proof of Theorem 1.6

We shall prove Theorem 1.6 by arguing that the second of the dichotomy cannot happen if $\{p \in M : K(p) < 0\}$ has only finitely many connected components.

Proof of Theorem 1.6. — Suppose $\Lambda \cap (\text{Per}\ (\Phi))^c \neq \emptyset$. Consider the two points y and z given by Theorem 1.4. We lift the geodesics $\gamma_y(t)$ and $\gamma_z(t)$ to the geodesics in the universal cover $\tilde{M}$, which are denoted by $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ respectively.

Consider the connected components of $\{p \in \tilde{M} : K(p) < 0\}$ on $\tilde{M}$ and we want to see how they distribute inside the ideal triangle bounded by $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$. Since $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are flat geodesics, any connected component doesn't intersect $\tilde{\gamma}_1$ or $\tilde{\gamma}_2$. We also claim that the radii of inscribed circles inside these connected components are bounded away from 0. Indeed, if we assume the contrary, then there exists an isometry between a connected component with very small radius of inscribed circle and a connected component of $\{p \in M : K(p) < 0\}$ on M. This is impossible because the number of the connected components of $\{p \in M : K(p) < 0\}$ is finite and the radii of their inscribed circles are bounded away from 0. The claim follows. Since $d(\tilde{\gamma}_1(t), \tilde{\gamma}_2(t)) \to 0$ as $t \to +\infty$, it is clear that the connected components of $\{p \in \tilde{M} : K(p) < 0\}$ cannot approach w inside of the ideal triangle. See Figure 4.

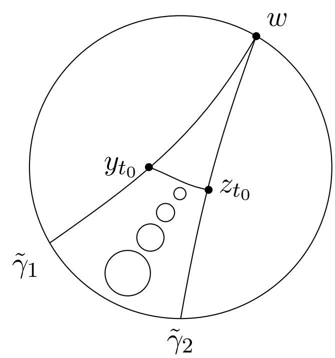

Figure 4. — Proof of Theorem 1.6

So there exist a $t_0 > 0$, $y_{t_0} = \Phi^{t_0}(y)$, $z_{t_0} = \Phi^{t_0}(z)$, such that the infinite triangle $z_{t_0} y_{t_0} w$ is a flat region. Then $d(\Phi^t(y), \Phi^t(z)) \equiv d(y_{t_0}, z_{t_0})$ for all $t \geqslant t_0$. Indeed, if we construct a geodesic variation between $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$, then Jacobi fields are constant for $t \geqslant t_0$ since $K \equiv 0$. Thus $d(\tilde{\gamma}_1(t), \tilde{\gamma}_2(t))$ is constant when $t \geqslant t_0$. We get a contradiction since $d(\Phi^t(y), \Phi^t(z)) \to 0$ as $t \to +\infty$ by Theorem 1.4.

Finally we can conclude that $\Lambda \subset \mathrm{Per}\ (\Phi)$. In particular the geodesic flow is ergodic by Theorem 1.3. $\square$

At last, let us suppose that $\{p \in M : K(p) < 0\}$ has infinitely many connected components. In the argument in the proof of Theorem 1.6, we cannot claim any more that the radii of inscribed circles inside connected components of $\{p \in \tilde{M} : K(p) < 0\}$ are bounded away from 0, as the radii of inscribed circles inside connected components of $\{p \in M : K(p) < 0\}$ could be arbitrarily small.

Question 4.1. — If $\{p \in M : K(p) < 0\}$ has infinitely many connected components, is it possible that $\lim_{t \to +\infty} d(\Phi^t(y), \Phi^t(z)) = 0$ for some $y, z \in \Lambda$, $y \notin \mathcal{O}(z)$?

A negative answer to Question 4.1 together with Theorem 1.3 will imply Conjecture 1.2, and in particular the ergodicity of the geodesic flow.

Acknowledgement. — The author would like to thank Federico Rodriguez Hertz for posing the problem and numerous discussions with him. The author would also like to express gratitude to Anatole Katok for his comments and constant encouragement. He also thanks Barbara Schapira, Keith Burns, Gerhard Knieper and the referee for valuable comments.

Bibliography

[1] Barreira (L.) and Pesin (Y.B.). — Lyapunov Exponents and Smooth Ergodic Theory, Vol. 23. AMS Bookstore (2002).

[2] Brin (M.) and Stuck (G.). — Introduction to dynamical systems, Cambridge University Press, (2002).

[3] Burns (K.) and Gelfert (K.). — Lyapunov spectrum for geodesic flows of rank 1 surfaces, Discrete Contin. Dyn. Syst. 34, no. 5, p. 1841-1872 (2014).

[4] Burns (K.) and Matveev (V. S.). — Open problems and questions about geodesics, arXiv preprint arXiv:1308.5417 (2013).

[5] Cao (J.) and Xavier (F.). — A closing lemma for flat strips in compact surfaces of non-positive curvature, (2008).

[6] Eberlein (P.). — When is a geodesic flow of Anosov type? I, Journal of Differential Geometry 8.3, p. 437-463 (1973).

[7] Eberlein (P.). — Geodesic flows in manifolds of nonpositive curvature, In Proceedings of Symposia in Pure Mathematics, vol. 69, p. 525-572. Providence, RI; American Mathematical Society; (1998), (2001).

[8] Eberlein (P.) and O'Neill (B.). — Visibility manifolds, Pacific Journal of Mathematics 46, no. 1, p. 45-109 (1973).

[9] Katok (A.) and Hasselblatt (B.). — Introduction to the modern theory of dynamical systems, Vol. 54. Cambridge university press (1997).

[10] Rodriguez Hertz (F.). — On the geodesic flow of surfaces of nonpositive curvature, arXiv preprint math/0301010 (2003).

[11] Ruggiero (R.). — Flatness of Gaussian curvature and area of ideal triangles, Bul. Braz. Math. Soc. 28, 1 p. 73-87 (1997).