Horocycle flow on negative variable curvature surface is standard

Adam Kanigowski  Kurt Vinhage*  Daren Wei†

Abstract

We provide a new proof that the horocycle flow preserving the Margulis measure on a variable negative curvature surface is standard. This was first proved by Ratner in [8]. The main purpose of this note is to provide a simplified case of the arguments in [4], which have similar but more complicated structures in the case of homogeneous flows on Lie groups, as well as illustrate the versatility of the method by applying it to a non-homogeneous flow.

1 Introduction

Let $M$ be a $C^3$ compact orientable surface of negative curvature and denote $U(M)$ as its unit tangent bundle. $U(M)$ carries a natural geodesic flow, $g_t$. It is well known that $g$ is a $C^2$ Anosov flow. As a result, we know for geodesic flow on $U(M)$, there exist invariant, orientable 1-dimensional foliations $\mathcal{H}$ and $\mathcal{K}$ of expanding curves and contracting curves, respectively. Since these foliations are 1-dimensional, it is natural to try to construct flows for which the leaves of $\mathcal{H}$ are orbits. There are many ways to do this, but two of them are the most natural: one may define $h_s(x)$ to be the point $y$ at distance $s$ from $x$ along the leaf $\mathcal{H}(x)$ (there are two such points, but the orientation on $\mathcal{H}(x)$ fixes the choice). Another is to do a similar process but with respect to a canonical measure on the leaf $\mathcal{H}(x)$, the disintegration of the Margulis measure.

We consider the second case in this paper, for a precise construction of the flow and its properties, see Section 3. Ergodic properties of such flows were studied by Marcus, Feldman and Ornstein, and Ratner, among others (see [6, 2, 8]). Marcus showed in [6] that such flows are uniquely ergodic. In [8], Ratner showed that these flows (in fact any such flow parameterizing the leaves of $\mathcal{H}(x)$) were standard, also called zero entropy loosely Bernoulli. We provide a new proof of this result using different methods:

Theorem 1.1. The horocycle flow $h_t$ preserving the Margulis measure is standard for any variable negative curvature surface.

The methods used here resemble those of a paper by the same authors [4], and this note may serve as a good starting point before the more technical and involved arguments found there.

*K. V. was supported by the National Science Foundation under Award DMS 1604796
†D. W. was partially supported by the NSF grant DMS-16-02409

1zero entropy loosely Bernoulli or loosely Kronecker
2 Preliminaries on Standard Systems

Standardness (zero entropy loosely Bernoulli or loosely Kronecker) is a concept introduced by A. Katok [5] and J. Feldman [1]. Standardness can be thought of as a measure of how closely a system resembles to a rigid action, as any standard system can be represented as a special flow over irrational rotation.

The notion of a standard system has many different equivalent definitions., as was shown in [5, 1, 9]. For our purposes, the most convenient is that of [9], we summarize the notions and definition here.

The basic setting is following: suppose \( h_t \) is an ergodic flow on a Lebesgue space \((X, \mathcal{B}, \mu)\) and \( \mathcal{P} \) is a finite measurable partition of \( X \). For \( x \in X \), denote \( \mathcal{P}(x) \) as the atom of \( \mathcal{P} \) which contains \( x \) and define \( I_R(x) = \{ T_s x : s \in [0, R] \} \) for \( R \geq 0 \). Let \( l \) as the Lebesgue measure on \([0, R]\).

**Definition 2.1 \((\varepsilon, \mathcal{P})\)-matchable, [9].** For \( x, y \in X, \varepsilon > 0 \) and \( R > 1 \), \( I_R(x) \) and \( I_R(y) \) are called \((\varepsilon, \mathcal{P})\)-matchable if there exists a subset \( A = A(x, y) \subseteq [0, R] \), \( l(A) > (1 - \varepsilon)R \) and an increasing absolutely continuous map \( \psi = \psi(x, y) \) from \( A \) onto \( A' = A'(x, y) \subseteq [0, R] \), \( l(A') > (1 - \varepsilon)R \) such that \( \mathcal{P}(T_s x) = \mathcal{P}(T_{h(t)} y) \) for all \( t \in A \) and derivative \( \psi' = \psi'(x, y) \) satisfies

\[
|\psi'(t) - 1| < \varepsilon \text{ for all } t \in A.
\]

we call \( \psi \) an \((\varepsilon, \mathcal{P})\)-matching from \( I_R(x) \) onto \( I_R(y) \).

Let

\[
f_R(x, y, \mathcal{P}) = \inf \{ \varepsilon > 0 : I_R(x) \text{ and } I_R(y) \text{ are } (\varepsilon, \mathcal{P})\text{-matchable} \}.
\]

While \( f_R \) is not a metric, it has enough similar properties to justify constructing \( f_R \)-balls. Let \( B_R(x, \varepsilon, \mathcal{P}) = \{ y \in X : f_R(x, y, \mathcal{P}) < \varepsilon \} \) be the \((R, \mathcal{P})\)-ball of radius \( \varepsilon > 0 \) centered at \( x \in X \), \( R > 1 \). \( B_R(x, \varepsilon, \mathcal{P}) \) should be thought of as analogous to a Hamming ball in entropy theory, with the exception that we allow for a small time change given by \( \psi \).

**Theorem 2.2** ([9], [10]). A zero-entropy ergodic measure preserving flow \( h_t \) is standard if and only if for every \( \varepsilon > 0 \) and a generating family of partitions \( \mathcal{P}_i \), there exists \( N(\varepsilon, \mathcal{P}_i) > 0 \) such that for every \( R > 0 \), there exist \( x_1, \ldots, x_n \) with \( n \leq N(\varepsilon, \mathcal{P}_i) \) such that:

\[
\mu \left( \bigcup_{k=1}^{n} B_R(x, \varepsilon, \mathcal{P}_i) \right) > 1 - \varepsilon
\]

Therefore, the standardness of the system can be detected by computing the decay rate (or lack thereof) of \( B_R(x, \varepsilon, \mathcal{P}) \) in \( R \). In fact, because if \( f_R(x, y, \mathcal{P}), f_R(y, z, \mathcal{P}) < \varepsilon, f_R(x, z, \mathcal{P}) < 5\varepsilon \), we know that a minimal cover of a set of size \( 1 - \varepsilon \) by \( \varepsilon\)-\((R, \mathcal{P})\) balls will have their \( \varepsilon/5\)-\((R, \mathcal{P})\) balls disjoint. In particular, if \( \mu(B_R(x, \varepsilon, \mathcal{P})) \) is bounded below independently of \( x \) and \( R \), we may conclude that \( h_t \) is standard.
3 Preliminaries on \( W^u \) flows

Let \( S \) be a compact, negatively curved, oriented surface, and \( g_s \) be the corresponding geodesic flow on its unit tangent bundle, \( M \). There exists a 1-dimensional unstable foliation, with smooth leaves \( W^u(x) \) for any \( x \in M \). Since \( S \) is oriented, the leaves \( W^u(x) \) can be given an orientation. We wish to define a continuous flow \( h_t \) whose orbits are exactly \( W^u(x) \), but such a flow will depend on the way we parameterize each leaf.

We will introduce a very special parameterization by first constructing the desired invariant measure:

3.1 Margulis measure

The Margulis measure for a transitive Anosov flow was introduced by G. A. Margulis in [7]. For any \( p \in M \), let \( W^s(p) \) be the leaf of the stable or unstable foliation through \( p \) for \( * = s, u \), respectively, and \( W^0(p) \) be the leaf of the center-stable or center-unstable foliation through \( p \) for \( * = s, u \), respectively. Recall that \( W^0s \) is the joint integration of the orbits of \( g^t \) and the \( W^s \) foliation, and similarly for \( W^0u \). Define \( OC(W^0u(p)) \) as the collection of open sets in \( W^0u(p) \) with compact closure and

\[
OC(W^0u) = \bigcup_{p \in M} OC(W^0u(p)).
\]

Let \( O \) be a subset of \( M \) contained in a sufficiently small neighborhood of a point \( p \). Then set \( f_O(x) = \mu^s((\{x\} \times U^s(p)) \cap O) \) where \( x \in U^0u(p) \) and \( U(p) \) is a neighborhood of \( p \) with local product form \( U(p) = U^0u(p) \times U^s(p) \) such that \( U^0u(p) \subset W^0u(p) \) and \( U^s(p) \subset W^s(p) \). Then the Margulis measure will be a finite \( g^t \)–invariant measure with the following form:

\[ \text{Proposition 3.1 (Margulis Measure, [7]).} \]

There exists a unique family of measures \( \mu^0u \) defined on the leaves \( W^0u(x), x \in M \), a unique family of measures \( \mu^s \) defined on the leaves \( W^s(x), x \in M \), and a unique \( g^t \)-invariant measure \( \mu \) on \( M \) such that:

(a) \( \mu^0u \) has the uniformly expansion property:

\[
\mu^0u(g^t(U)) = e^{h^u t} \mu^0u(U)
\]

where \( U \subset OC(W^0u), h^u > 0 \) and \( t \in \mathbb{R} \).

(b) \( \mu^s \) has the uniformly contraction property:

\[
\mu^s(g^t(U')) = e^{h^u t} \mu^s(U')
\]

where \( U' \subset OC(W^s), h^s < 0 \) and \( t \in \mathbb{R} \).

(c) Letting \( \mu_q(A) = \mu^0u(A \times \{q\}) \) for \( q \in U^s(p), A \subset U^0u(p) \), we have

\[
\mu(O) = \int f_O(x) d\mu_q(x).
\]

Furthermore, \( \mu \) is the unique measure of maximal entropy for \( g^t \)
Remark 3.2. Because $\mu$ is the unique measure of maximal entropy for $g^t$, it is also the unique measure of maximal entropy for $g^{-t}$, another Anosov flow. Therefore, $\mu$ disintegrates in the analogous way for $\mu^0_s$ and $\mu^u$.

We now fix the flow $h_t$ whose orbits are the leaves $W^u(x)$: if $x \in M$, let $h_t(x)$ be the point $y$ such that the interval connecting $x$ and $y$ has $\mu^u$-measure exactly $t$. Of course, there are two such points, but the leaves $W^u(x)$ are oriented, so if $t > 0$, we choose $y$ in the positive direction, and if $t < 0$, we choose $y$ in the negative direction.

Note that in most cases, this action is only Hölder, and is not even generated by a vector field. The ergodic properties of these flows were first studied by B. Marcus in [6], and later by Feldman and Ornstein in [2]. We remark that making this choice for $h_t$ yields

$$g_s h_t = h_{e^t g_s}$$

by Proposition 3.1(b) and Remark 3.2. Notice also that by a completely symmetric argument, an analogous flow $k_t$ may be built whose orbits are the leaves of $W^s$, and

$$g_s k_t = h_{e^{-t} g_s}.$$

Finally, because of the product structure provided by Proposition 3.1(c), the map:

$$f : (r,s,t) \mapsto g_r h_s k_t (x)$$

is a homeomorphism onto a neighborhood of $x$, and $f_s (dr \times ds \times dt) = \mu$.

3.2 Estimates relating the flows $h_t$, $k_t$ and $g_t$

In this section, we recall some lemmas from [2] which will be helpful for the proof of the main theorem. We remark that for clarity, we use $h_t$ and $k_t$ to denote the flows coming from the Margulis parameterization as described in the previous subsection, while the notation of [2] uses $\bar{h}_t$ and $\bar{k}_t$ to denote these flows.

Since Anosov flows have a local product structure for their weak-stable and unstable manifolds, if $y \in W^u(x)$ and $z \in W^{0s}(x)$ are sufficiently close to $x$, there is a unique point in $W^{0s}_{loc}(y) \cap W^{u}_{loc}(z)$. Hence, there are unique continuous functions $\rho(x,s,t), \sigma(x,s,t), \tau(x,s,t)$ such that

$$g_u k_r h_s x = h_{\sigma} k_t x,$$

(4)

$\sigma$ and $\tau$ has the same sign as $s,t$, respectively, and $\rho(x,0,t) = \rho(x,s,0) = 0$. Furthermore, $\rho(x,s,t)$ enjoys the equivariance property:

$$\rho(x, e^{-u} s, e^{u} t) = \rho(g_u x, s, t)$$

(5)

In particular, we may define $\rho$ not only for sufficiently small $s,t$, but for all $s,t$ such that $|st|$ is sufficiently small. Similarly, we get definitions on such $s,t$ for $\sigma, \tau$. We have the following results of [2], the first of which appeared as Lemmas 3.4 and 2.8, respectively:

Lemma 3.3. The function $\sigma$ is differentiable in the second argument, and for sufficiently small $st$,

$$\frac{\partial \sigma}{\partial s}(x,s,t) = e^{\rho(x,s,t)}.$$
Lemma 3.4. Given $\varepsilon > 0$, $\alpha$ may be chosen so small that if $y = g_u h_x k_w x$ with $|u|, |v|, |w| < \alpha$ and $|s| < \frac{\alpha^2}{2|w|}$, then $d(h_x x, h_{e^{u\sigma(x,s,w)}} y) < \varepsilon$.

3.3 $u$-partitions

Recall that when we consider the standardness, it is very crucial to consider the suitable generating partition. In this paper, we will use the idea of $u$−partition which was introduced in [8].

Definition 3.5 ($u$−isomorphic and $s$−isomorphic [8]). Two sets $A, B \subset W^0s$ are called $u$−isomorphic if there is continuous $\psi : A \times I \to M$ s.t.

- $\psi(x, I) \subset W^u$, $x \in A$;
- $\psi(x, 0) = x$, $\psi(x, 1) \in B$ and the map $\tilde{\psi} : A \to B$, $\tilde{\psi}(x) = \psi(x, 1)$ is a homeomorphism.

The set $\psi(A \times I) = P$ is called a $u$−cylinder with faces $A, B$. If the positive direction on orbits of $h_t$ goes from $A$ to $B$ we write $A = A_1(P)$, $B = A_2(P)$. $\psi(x, I)$ and $\psi(y, I)$, $x, y \in A$ are called $s$−isomorphic.

Definition 3.6 ($u$−partition [8]). Let $\alpha = \{P_1, \ldots, P_m\}$ be a partition of $M$ into $u$−cylinders and we will say $\alpha$ is a $u$−partition.

Recall that Margulis measure has local product structure, thus the boundary of this partition has measure zero. Another important feature of this partition is that we can formulate a family of generating partitions based on this partition. In fact, by shrinking the size of $A, B$ and distance of $A, B$, it is clear that we can construct a family of generating partitions:

Lemma 3.7 (Ratner, [8]). For every $\delta > 0$, there exists a $u$−partition of $M$ such that $\text{diam}(P_i) < \delta$

4 Proof of Theorem 1.1

In this section, we will prove the Theorem 1.1. Before we give a proof of this theorem, we will introduce some definitions and tools which will help us formulate the proof.

From now, we will prove the Theorem 1.1 for a fixed $u$−partition. And then the same proof can be applied to any other partitions and thus we done.

Fix a $u$−partition $\alpha_m$, let $V^{m}_{\varepsilon^2}$ be the $\varepsilon^2$ neighborhood of the boundary of $\alpha_m$. Since Margulis measure $\mu$ has a local product measure on $W^0s$ and $W^u$, it follows that $\mu(V^{m}_{\varepsilon^2}) = O(m^2\varepsilon^2)$. By ergodic theorem on $h_t$ and $\chi_{V^{m}_{\varepsilon^2}}$, we obtain a set $D_{\varepsilon}$ such that $\mu(D_{\varepsilon}) > 1 - \varepsilon^2$ and a number $N_{\varepsilon} > 0$ such that for every $R \geq N_{\varepsilon}$ and every $y \in D_{\varepsilon}$, we have

$$|\{t \in [0, R] : h_t(y) \in V^{m}_{\varepsilon^2}\}| \leq \frac{\varepsilon R}{2}. (6)$$

If we can prove the upper bound of the number of Kakutani matching balls is bounded, then we finish our proof. Indeed, the upper bound of number of Kakutani balls will follow from the following three claims:
Definition 4.1 (Pre-Matching Balls). If $x, y \in M$ we say that $x \in PM(R, \varepsilon, y)$ if and only if $x = g_u h_v k_w y$ and $|w| < \varepsilon/R$, $|u| < \varepsilon$ and $|v| < \varepsilon$.

Notice that since the coordinates $(u, v, w) \rightarrow g_u h_v k_w y$ take Lebesgue measure to the Margulis measure, $\mu(PM(R, \varepsilon, y)) = \varepsilon^3/R$ (recall the end of Section 3.3).

Claim 4.2. For every $y \in D_\varepsilon$, every $R \geq 1$ and every $x \in M$ of there exists $p \in [0, \varepsilon^3 R]$ such that if for $|u|, |v| < \varepsilon^5$ and $|w| < \frac{\varepsilon^5}{5}$,
\[ h_{p} x \in PM(R, \varepsilon^5, y) \]

Then $x \in B_R(y, \varepsilon, \alpha_m)$.

Claim 4.3. For every $y \in D_\varepsilon$ and every $p, q \in [0, \varepsilon^3 R]$, with $|p - q| \geq 1$, we have
\[ h_{-p}(PM(R, \varepsilon^5, y)) \cap h_{-q}(PM(R, \varepsilon^5, y)) = \emptyset. \]

Let us prove Theorem 1.1 before showing each claim:

Proof of Theorem 1.1. Suppose $y \in D_\varepsilon$. Then by claim 4.2 we have
\[ \bigcup_{p \in [0, \varepsilon^3 R]} h_{-p}(PM(R, \varepsilon^5, y)) \subset B_R(y, \varepsilon, \alpha_m). \]

By claim 4.3 and the formula for $\mu(PM(R, \varepsilon^5, y))$, we have
\[ \mu(B_R(y, \varepsilon, \alpha_m)) \geq \mu\left( \bigcup_{p \in [0, \varepsilon^3 R]} h_{-p}(PM(R, \varepsilon^5, y)) \right) \geq \varepsilon^3 R \mu(PM(R, \varepsilon^5, y)) \geq \varepsilon^{18}. \]

By the remarks following the statement of Theorem 2.2 we conclude $h_t$ is standard.

Proof of Claim 4.2. Take $y \in D_\varepsilon$ and let $p \in [0, \varepsilon^3 R]$ be such that $h_p x \in PM(R, \varepsilon^5, y)$.

By Definition 4.1 we have for some $|u|, |v| < \varepsilon^5$, $|w| < \frac{\varepsilon^5}{5}$ such that
\[ x = h_{-p} g_u h_v k_w y. \quad (7) \]

Let $\psi(t) = e^{u} \sigma(x, t, w)$ as Lemma 3.1, $l(t) = \psi(t) + p$, $A(x, y) = \{ t \in [0, R] : h_t(y) \notin V_{\varepsilon^2} \}$. Moreover, by Lemma 3.3 and (5):
\[ l'(t) = e^{u} \partial_t \sigma(x, t, w) = e^{u + \rho(x, t, w)} = \exp(u + \rho(g_{-\log t} x, \varepsilon, \omega t/w)). \]

Notice that $u$ is small and $|w| < \varepsilon^5/R < \varepsilon^5/t$, so $|wt/\varepsilon| < \varepsilon^4$. Therefore, since $\rho(x, \varepsilon, 0) = 0$ for all $x$ and $(x, t) \rightarrow \rho(x, \varepsilon, t)$ is jointly continuous, we have, for sufficiently small $\varepsilon$ depending on the fixed function $\rho$,
\[ |l'(t) - 1| < \varepsilon \]

for every $t \in [0, R]$ and hence $l$ satisfies the matching function condition of $(\varepsilon, \alpha_m)$.

By Lemma 3.4 we know that
\[ d(h_t y, h_{l(t)} x) \leq \varepsilon. \quad (8) \]

Thus for every $t \in A(x, y)$, we have $\alpha_m(h_t y) = \alpha_m(h_{l(t)} x)$. Since $|A(x, y)| \geq (1 - \varepsilon) R$, thus we have $x \in B_R(y, \varepsilon, \alpha_m)$ and finish the proof. 

\[ \square \]
Proof of Claim 4.3: Suppose that there exists $x \in h^{-q}(PM(R, \varepsilon^5, y)) \cap PM(R, \varepsilon^5, y)$ with $\varepsilon^3 R \geq |p - q| \geq 1$ and $y \in D_x$. Thus by the Definition 4.1 and denote $r = p - q$, we have

$$x = g_u h_v k_w y$$

and

$$h_{-r} x = g_{u'} h_{v'} k_{w'} y$$

where $|u|, |u'|, |v|, |v'| \leq \varepsilon^5$ and $|w|, |w'| \leq \frac{\varepsilon}{R}$.

Therefore using the first equality to express $x$ in the second equality, we have,

$$h_{-r} g_u h_v k_w y = g_{u'} h_{v'} k_{w'} y.$$

Recall that $h_{-r} g_u = g_u h_{-e^{-u}r}$ thus we have

$$g_u h_{-e^{-u}r} k_w y = g_{u'} h_{v'} k_{w'} y. \quad (9)$$

Then acting by $k_{-w'} h_{-v'} g_{-w'}$ on the both sides of (9), we have

$$k_{-w'} h_{-v'} g_{-w'} h_{-u'} k_{w'} y = y. \quad (10)$$

We now further act on both sides of (10) by $g_{-t}$ and pushing it through to $y$ via conjugation. We select $t = \log(\varepsilon^{-2} r)$, to get

$$(g_{-t} k_{-w'} g_t)(g_{-t} h_{-v'} g_{t} g_{u'}(g_{-t} h_{-u'} g_{t} g_{e^{-t} r} g_{t}) g_{-t} y = g_{-t} y. \quad (11)$$

By renormalization relation, we have,

$$k_{-e^{-t} w'} h_{-e^{-t} v'} g_{u'} h_{e^{-t} (v - e^{-u} r)} k_{e^{-t} w} g_{-t} y = g_{-t} y. \quad (12)$$

Since $\max\{|e^{-t} w'|, |e^{-t} v'|, |u - w'|, |e^t w|\} \leq \varepsilon^5$ and $2 \varepsilon \geq |e^{-t} (v - e^{-u} r)| \geq \varepsilon^2$ then we have the left side of (12) is $\varepsilon^2$ away from $g_{-t} y$ and thus we get a contradiction. \qed

References

[1] J. Feldman, New K-automorphisms and a problem of Kakutani, Israel J. Math. 24.1 (1976): 16-38.

[2] J. Feldman, D. Ornstein, Semirigidity of horocycle flows over compact surfaces of variable negative curvature. Ergodic Theory Dynam. Systems 7 (1987), no. 1, 49–72.

[3] V. A. Kaimanovich, Bowen-Margulis and Patterson measures on negatively curved compact manifolds. Dynamical systems and related topics (Nagoya, 1990), 223–232, Adv. Ser. Dynam. Systems, 9, World Sci. Publ., River Edge, NJ, 1991.

[4] A. Kanigowski, K. Vinhage, D. Wei, Kakutani equivalence of unipotent flows, preprint, arXiv:1805.01501.

[5] A. B. Katok, Time change, monotone equivalence, and standard dynamical systems. (Russian) Dokl. Akad. Nauk SSSR 223 (1975), no. 4, 789–792.
[6] B. Marcus, *Ergodic properties of horocycle flows for surfaces of negative curvature*. Ann. of Math. (2) 105 (1977), no. 1, 81–105.

[7] G. A. Margulis, *Certain measures that are connected with U-flows on compact manifolds*. (Russian) Funkcional. Anal. i Priložen. 4 1970 no. 1, 62–76.

[8] M. Ratner, *Horocycle flows are loosely Bernoulli*. Israel J. Math. (1978), 31: 122–132.

[9] M. Ratner, *Some invariants of Kakutani equivalence*, Israel J. Math. 38 (1981), no. 3, 231–240.

[10] M. Ratner, *Time change invariants for measure preserving flows*. (English summary) Modern theory of dynamical systems, 263–273, Contemp. Math., 692, Amer. Math. Soc., Providence, RI, 2017.