Ergod. Th. &. Dynam. Sys. (1982), **2**, 339–367
Printed in Great Britain

Entropy and closed geodesics

A. KATOK
Department of Mathematics, University of Maryland, College Park, Maryland 20742, *USA*

(*Received* 3 *April* 1981)

To the memory of my friend Volodya Alexeyev (1932–1980)

Abstract. We study asymptotic growth of closed geodesics for various Riemannian metrics on a compact manifold which carries a metric of negative sectional curvature. Our approach makes use of both variational and dynamical description of geodesics and can be described as an asymptotic version of length–area method. We also obtain various inequalities between topological and measure-theoretic entropies of the geodesic flows for different metrics on the same manifold. Our method works especially well for any metric conformally equivalent to a metric of constant negative curvature. For a surface with negative Euler characteristics every Riemannian metric has this property due to a classical regularization theorem. This allows us to prove that every metric of non-constant curvature has strictly more close geodesics of length at most T for sufficiently large T then any metric of constant curvature of the same total area. In addition the common value of topological and measure-theoretic entropies for metrics of constant negative curvature with the fixed area separates the values of two entropies for other metrics with the same area.

1. *Introduction*

1(A). *Notations.* Throughout this paper M will always denote a smooth, compact connected manifold without boundary.

Let σ be a Riemannian metric on M (normally of class C^2, unless the opposite is stated). The metric σ generates a volume element on M (Riemannian volume form). We will use the following standard notations for various objects associated with σ:

d_σ – the distance function generated by σ;

v_σ – the total volume (area if $\dim M = 2$) of M;

μ_σ – normalized Riemannian measure on M which assigns to every set its volume divided by v_σ;

$l_\sigma(\alpha)$ – the length of a curve α in the metric generated by σ;

$S^\sigma M$ – the unit tangent bundle over M, i.e. the manifold of all tangent vectors to M of length one; there is a canonical way to define a Riemannian metric on $S^\sigma M$;

D_σ – the distance in $S^\sigma M$ generated by this metric;

λ_σ – the normalized Riemannian measure on $S^\sigma M$; it is sometimes called the Liouville measure;

$\pi_\sigma: S^\sigma M \to M$ – the standard projection. It is easy to see that $(\pi_\sigma)_*\lambda_\sigma = \mu_\sigma$;

$\phi^\sigma = \{\phi_t^\sigma\}_{t\in\mathbb{R}}$ – the geodesic flow generated by the metric σ, i.e. a one-parameter group of diffeomorphisms of $S^\sigma M$ determined by the motion of tangent vectors with unit speed along geodesics determined by them.

The flow ϕ^σ preserves the measure λ^σ. Our standard assumption guarantees that ϕ^σ is a C^1 flow;

$D_\sigma^t(v_1, v_2) = \max_{0\le s\le t} D_\sigma(\phi_s^\sigma v_1, \phi_s^\sigma v_2)$;

h_σ – the topological entropy of the geodesic flow ϕ_t^σ;

$h_0^\nu = h_\nu(\phi^\sigma)$ – the metric entropy of ϕ^σ with respect to an invariant measure ν. For brevity we will write h_σ^λ instead of $h_\sigma^{\lambda^\sigma}$.

If it does not lead to a confusion we will not make any distinction between a geodesic (finite or infinite) as a curve in M parametrized by length and its lifting to the unit tangent bundle which is an orbit of the geodesic flow. We will call both objects geodesics and will normally use the symbol γ with various indices to denote them.

By a closed geodesic we will normally mean a simple periodic orbit of the geodesic flow. Thus, every simple closed geodesic, as it is usually defined in differential geometry, determines two closed geodesics in our sense: one for each orientation.

Let us introduce a few more notations.

$\Pi(M)$ – the set of all non-zero free homotopy classes of closed curves on M;

For $\Gamma \in \Pi(M)$ we denote

$$L_\sigma(\Gamma) = \inf_{\alpha\in\Gamma} l_\sigma(\alpha);$$

$\mathcal{P}_\sigma(T)$ – the set of all closed geodesics of length $\le T$;

$P_\sigma(T) = \operatorname{Card} \mathcal{P}_\sigma(T)$ – the number of closed geodesics of length less than or equal to T; (1.1)

$P_\sigma^s(T)$ – the number of $\Gamma \in \Pi$ such that $L_\sigma(\Gamma) \le T$;

$$P_\sigma = \varlimsup_{T\to\infty} (\log P_\sigma(T)/T); \qquad P_\sigma^s = \varlimsup_{T\to\infty} (\log P_\sigma^s(T)/T). \tag{1.2}$$

Let σ_1, σ_2 be two Riemannian metrics on M and $v \in S^{\sigma_1}M$ then $\|v\|_{\sigma_2}$ is the norm of v with respect to σ_2,

$$[\sigma_1; \sigma_2] = \int_{S^{\sigma_1}M} \|v\|_{\sigma_2} d\lambda_{\sigma_1}(v).$$

1(B). *Main results.* We begin in § 2 with the proof of the following basic theorem.

THEOREM A. *Let σ_1 be a Riemannian metric of negative curvature. Then for every Riemannian metric σ_2 on the same manifold*

$$P_{\sigma_2}^s \ge ([\sigma_1; \sigma_2])^{-1} h_{\sigma_1}^\lambda.$$

Then we prove a weaker inequality with h_{σ_2} instead of $P_{\sigma_2}^s$ under a weaker assumption that σ_1 is a metric without focal points. We finish the section with a lengthy discussion of the nature of the extra multiples involved in the inequalities which allows us to sharpen the inequalities under some additional assumptions.

In § 3 the results of the previous section are applied to the two-dimensional case producing the main result of this paper.

THEOREM B. *For every Riemannian metric* σ *on a compact surface with negative Euler characteristic* E

$$P_\sigma \geq h_\sigma \geq P_\sigma^s \geq (-2\pi E/v_\sigma)^{\frac{1}{2}}$$

and the last inequality is strict for every metric of non-constant curvature.

If σ *is a metric without focal points then*

$$(-2\pi E/v_\sigma)^{\frac{1}{2}} \geq h_\sigma^\lambda$$

and this inequality is strict unless σ *is a metric of constant negative curvature.*

Then we show how these inequalities can be sharpened according to various geometric characteristics of σ. In § 4 we consider several refinements of theorem B.

1(C). *Background.* In order to simplify the references in the course of proofs and at the same time give the reader an idea of our approach we list here the principal facts used in the proofs of theorems A and B. Some of these facts are classical, others are fairly recent and relatively less known.

1.1. *Every free homotopy class* $\Gamma \in \Pi(M)$ *contains a shortest curve and every such curve is a closed geodesic* (cf. e.g. [**5**, § 11.7]).

1.2. *For a metric of negative (sectional) curvature on a compact manifold, a closed geodesic in every non-trivial free homotopy class is unique.*

1.3. *Anosov closing lemma* [**1**]. *If* σ *is a metric of negative curvature, then for every* $\varepsilon > 0$ *there exists* $\delta = \delta(\varepsilon, \sigma) > 0$ *such that if for* $v \in S^\sigma M$, $t > 0$

$$D_\sigma(v, \phi_t^\sigma v) < \delta,$$

then there exists a closed geodesic $\gamma_v = \{\phi_s^\sigma w\}_{s=0}^{t'}$ *of period* t', *where* $|t - t'| < \varepsilon$, *such that*

$$D_\sigma(\phi_s^\sigma v, \phi_s^\sigma w) < \varepsilon \quad \text{for } 0 \leq s \leq t.$$

1.4. *For a metric of negative curvature, the limit of* $(\log P_\sigma(T))/T$ *exists and is equal to* h_σ [**6, 19, 20**].

1.5. $P_\sigma^s \leq h_\sigma$.

Since P_σ^s is less or equal than the speed of exponential growth of balls on the universal covering (see 2.6, below), this inequality follows from Manning [**17**]. It is implicit already in Dinaburg [**8**].

The next proposition gives a characterization of the metric entropy for an ergodic flow through the growth of distinguishable typical orbits, which is an exact analogue and an easy corollary of the characterization given in [**13**, theorem 1.1] for discrete-time dynamical systems.

Let X be a compact metric space with distance function d, $f=\{f_t\}_{t\in\mathbf{R}}$ is a continuous flow on X, i.e. one-parameter group of homeomorphisms of X, μ is a Borel probability measure on X invariant and ergodic with respect to f. Let us define for every $T>0$ a metric d_T^f by

$$d_T^f(x,y)=\max_{0\le\tau\le T} d(f_\tau x, f_\tau y). \tag{1.3}$$

(Obviously, the metric D_σ^T defined above coincides with d_T^f for the geodesic flow ϕ^σ if D_σ serves as d.) Furthermore, let, for $T>0$, $\varepsilon>0$, $0<\delta<1$, $N_f^\mu(T,\varepsilon,\delta)$ be the minimal number of ε-balls in the metric d_T^f which covers a set of measure $\ge 1-\delta$. For the geodesic flow ϕ^σ we will write $N_{\phi_\sigma}(T,\varepsilon,\delta)$ instead of $N_{\phi_\sigma}^{\lambda_\sigma}(T,\varepsilon,\delta)$.

1.6. PROPOSITION. *For every* δ, $0<\delta<1$

$$h_\mu(f)=\lim_{\varepsilon\to 0}\varliminf_{T\to\infty}\frac{\log N_f^\mu(T,\varepsilon,\delta)}{T}=\lim_{\varepsilon\to 0}\varlimsup_{T\to\infty}\frac{\log N_f^\mu(T,\varepsilon,\delta)}{T}.$$

Proof. Obviously for every $t>0$

$$d_T^f(x,y)\ge d_{[T/t]}^{f_t}(x,y)\stackrel{\text{def}}{=}\max_{0\le k\le[T/t]} d(f_{kt}x, f_{kt}y).$$

On the other hand, by the uniform continuity of the flow, one can find for every $\varepsilon>0$, $t>0$ an $\varepsilon_1=\varepsilon_1(\varepsilon,t)>0$ such that $d(x,y)<\varepsilon_1$ implies that $d(f_\tau x, f_\tau y)<\varepsilon$ for $0\le\tau\le t$. This means that any ε-ball in d_T^f metric contains an ε_1-ball in $d_{[T/\tau]}^f$ metric for any $0\le\tau\le t$. Let, as in [**13**], $N_{f_t}^\mu(n,\varepsilon,\delta)$ be the minimal number of ε-balls in $d_n^{f_t}$-metric covering a set of measure $\ge 1-\delta$. We have from the above arguments

$$N_{f_t}^\mu\left(\left[\frac{T}{t}\right],\varepsilon_1,\delta\right)\ge N_f^\mu(T,\varepsilon,\delta)\ge N_{f_t}^\mu\left(\left[\frac{T}{t}\right],\varepsilon,\delta\right). \tag{1.4}$$

It is well-known that

$$h_\mu(f_t)=|t|h_\mu(f). \tag{1.5}$$

Let us choose $t>0$ such that f_t is ergodic so that we can apply theorem 1.1 from [**13**] to f_t. Now the proposition follows immediately from (1.4) and (1.5). □

1.7. *Any Riemannian metric σ on a compact surface with negative Euler characteristic can be represented in a unique way as $\rho\sigma_0$ where ρ is a positive scalar function and σ_0 is a metric of constant negative curvature such that $v_{\sigma_0}=v_\sigma$.*

This follows from the classical regularization theorem by Koebe (cf. e.g. [**27**]).

1.8. *For any $C^{2+\delta}$ ($\delta>0$) Riemannian metric σ on a compact surface*

$$P_\sigma\ge h_\sigma.$$

This follows from the continuous-time version of theorem 4.3 from [**13**]. We postpone a more detailed explanation till § 4 where we will discuss a more refined version and generalizations of that result.

1(D). *Historical remarks.* Geodesics on a Riemannian manifold can be studied from two different viewpoints, variational and dynamical. In differential geometry the first approach prevails, so that the geodesics are usually treated as shortest (locally or globally) elements in various classes of curves on the manifold. Variational method in the study of closed geodesics goes back to D. Hilbert and G. D. Birkhoff. The book of M. Morse [**21**] is a classical example of variational approach.

Morse was also the first to consider the asymptotic behaviour of geodesics for various metrics on the same manifold [**22**]. More specifically he showed that certain features of the picture corresponding to a metric of constant negative curvature remain for an arbitrary metric on a surface of negative Euler characteristic.

The length–area method known for a long time in complex analysis (see e.g. [**12**]) was first used in the problem of closed geodesics by Loewner (see Berger's lectures [**2**]).

From the dynamical point of view the geodesics are considered as orbits of the geodesic flow and the geometric characteristics of the metric are reflected in asymptotic properties of this flow; e.g. the negativity of the curvature leads to exponential divergence of the orbits [**1**].

Geodesic flows served as a standard proving ground for the modern global theory of smooth dynamical systems since its origin in works of H. Poincaré [**24**] and G. D. Birkhoff [**4**]. During the twenties and the thirties, geodesic flows especially on manifolds of negative curvature were studied by E. Artin, M. Morse, E. Hopf, G. A. Hedlund and others (cf. the survey of Hedlund [**10**] which summarizes most of the achievements of that period). These authors were mostly concerned with qualitative questions like the existence of closed geodesics, ergodicity, density of closed geodesics, etc.

Selberg [**28**] found the asymptotic growth rate for the number of closed geodesics on a manifold of constant negative curvature, based on his celebrated trace formula.

In the early sixties important progress was made following the introduction of new fundamental dynamical concepts – Anosov systems [**1**], measure-theoretic entropy (Kolmogorov–Sinai), and topological entropy (Adler–Konheim–McAndrew). Sinai [**29**] proved that for a Riemannian metric on an n-dimensional manifold with negative sectional curvature bounded between $-K_1^2$ and $-K_2^2$ $(K_1 < K_2)$

$$(n-1)K_1 \le P_\sigma \le (n-1)K_2.$$

Dinaburg [**8**] and Manning [**17**] found important connections between the topological entropy and the growth of closed geodesics. Margulis ([**19, 20**], cf. also [**6**]) proved that for a metric of negative curvature $P_\sigma(T)$ is multiplicatively equivalent to $(C \exp h_\sigma T)/T$ for a certain constant $C > 0$. Recently C. Toll (unpublished) showed that $C = h_\sigma^{-1}$.

1(E). I was working on this paper for almost two years. It was started during my visit to the Mathematics Institute of the University of Warwick in May–June 1979. I would like to thank the Institute, and to thank personally the organizer of the symposium of Diffeomorphisms and Foliations, D. Epstein, for their warm hospitality. Conversations with D. Epstein, A. Manning and A. Douady were extremely useful and stimulating for the early stages of my work.

Substantial progress was reached during my short visit to SFB 'Theoretische Mathematik' of the University of Bonn in June 1980. I would like to express my gratitude for this invitation. Discussion with differential geometers there, especially with H. Karcher, contributed substantially to this paper.

I have discussed various aspects of this paper with my colleagues and visitors at the University of Maryland. I am thankful to V. Ballmann, M. Berger, M. Brin, A. Gray, H. Gluck, B. Reinhart, R. Spatzier, S. Wolpert, and W. Ziller for their help.

Some of the results included in this paper are announced in the preprints of my talks at the 'Mathematische Arbeitstagung 1980' [**14**] and at the conference on Ergodic Theory in Durham, England, [**15**].

2. *Comparison theorems*

2(A). THEOREM 2.1. *Let σ_1 be a metric of negative curvature, on a compact manifold M, ν be a Borel probability measure invariant with respect to the geodesic flow φ^{σ_1}, σ_2 be another Riemannian metric on M and $\nu_{\sigma_2} = \int_{S^{\sigma_1}} \|v\|_{\sigma_2} d\nu$. Then*

$$P_{\sigma_2} \geq h^{\nu}_{\sigma_1}/\nu_{\sigma_2}.$$

Proof. First, let us notice that it is enough to prove the inequality only for an ergodic measure ν. For, otherwise ν can be decomposed into ergodic components. Both the entropy $h^{\nu}_{\sigma_1}$ and the integral ν_{σ_2} are the integrals of the corresponding quantities for ergodic measures. Since for any positive function g and non-negative function f on any measure space

$$\int f d\mu \Big/ \int g d\mu \leq \sup (f/g) \tag{2.1}$$

we see that the statement of the theorem holds for arbitrary ν if it holds for ergodic measures.

Let us denote for $T>0$, $\varepsilon>0$

$$A_{\varepsilon,T} = \left\{ v \in S^{\sigma_1}M : \left| \frac{1}{T} \int_0^T \|\phi_t^{\sigma_1} v\|_{\sigma_2} \, dt - \nu_{\sigma_2} \right| < \varepsilon \right\}. \tag{2.2}$$

$$B_{\varepsilon,T} = \{ v \in S^{\sigma_1}M : \exists t,\ T \leq t \leq (1+\varepsilon)T;\ D_{\sigma_1}(v, \phi_t^{\sigma_1} v) < \varepsilon \}.$$

Since the geodesic flow is ergodic with respect to ν then for any $\varepsilon>0$

$$\nu(A_{\varepsilon,T}) \to 1 \text{ as } T \to \infty. \tag{2.3}$$

Let us show that also for each $\varepsilon>0$

$$\nu(B_{\varepsilon,T}) \to 1 \quad \text{as } T \to \infty. \tag{2.4}$$

For, let ξ be a finite partition of $S^{\sigma_1}M$ into measurable sets of diameter less than ε. Let $c \in \xi$. Since by the ergodicity of ϕ^{σ_1} for almost every $v \in c$ we have

$$\lim \frac{1}{T} \int_0^T \chi_c(\phi_t^{\sigma_1} v) \, dt = \nu(c),$$

(χ_c is the characteristic function of the set c) then for sufficiently big T

$$\int_0^T \chi_c(\phi_t^{\sigma_1}v)\,dt < \int_0^{(1+\varepsilon)T} \chi_c(\varphi_t^{\sigma_1}v)\,dt$$

so that $\varphi_{t'}^{\sigma_1}v \in c$ for some t': $t<t'\le(1+\varepsilon)t$. Applying this argument to every c we obtain (2.4).

Thus by (2.3) and (2.4) we can find T such that for every $t>T$

$$\nu(A_{\varepsilon,t}\cap B_{\delta(\varepsilon,\sigma_1),t})>\tfrac{1}{2} \tag{2.5}$$

where $\delta(\varepsilon,\sigma_1)$ comes from 1.3.

Let $\Lambda_{t,\varepsilon}$ be a maximal $(t,3\varepsilon)$-separated subset of the set $A_{\varepsilon,t}\cap B_{\delta(\varepsilon,\sigma_1),t}=A$. The balls of radius 3ε about the points of the set $\Lambda_{t,\varepsilon}$ cover the set A (otherwise $\Lambda_{t,\varepsilon}$ is not maximal). Thus, we have by (2.5)

$$\text{Card}\,\Lambda_{t,\varepsilon}\ge N_{\varphi^{\sigma_1}}(t,3\varepsilon,\mu(A))\ge N_{\phi^{\sigma_1}}(t,3\varepsilon,\tfrac{1}{2}). \tag{2.6}$$

Applying Anosov's closing lemma 1.3, we construct for every $v\in\Lambda_{t,\varepsilon}$ a closed geodesic γ_v. Since by this lemma γ_v contains an element $w=w(v)$ such that $D^t_{\sigma_1}(v,w(v))<\varepsilon$ we have for $v_1\ne v_2$

$$\begin{aligned}D^t_{\sigma_1}(w(v_1),w(v_2))&>D^t_{\sigma_1}(v_1,v_2)-D^t_{\sigma_1}(v_1,w(v_1))-D^t_{\sigma_1}(v_2,w(v_2))\\&>3\varepsilon-\varepsilon-\varepsilon=\varepsilon.\end{aligned}$$

Thus, if the same closed geodesic appears as γ_{v_1} and γ_{v_2} the points $w(v_1)$ and $w(v_2)$ lie more than ε apart. Since the maximal number of such points on a geodesic does not exceed its length divided over ε we see that the total number of different geodesics appearing as γ_v for different $v\in\Lambda_{t,\varepsilon}$ is greater than

$$\frac{\varepsilon\cdot\text{Card}\,\Lambda_{t,\varepsilon}}{t(1+\varepsilon)}\ge\frac{\varepsilon}{t(1+\varepsilon)}N_{\phi^{\sigma_1}}(t,3\varepsilon,\tfrac{1}{2}). \tag{2.7}$$

Since σ_1 has negative curvature, all these geodesics belong to different free homotopy classes (see 1.2).

We want to estimate from above the σ_2-length of the geodesic γ_v. The number $t'<(1+\varepsilon)t$ is a period of this geodesic (probably not the minimal period) so that we have for some natural number n

$$\begin{aligned}l_{\sigma_2}(\gamma_v)&=\int_0^{t'/n}\|\phi_s^{\sigma_1}w(v)\|_{\sigma_2}\,ds\le\int_0^{t'}\|\phi_s^{\sigma_1}(w(v))\|_{\sigma_2}\,ds\\&=\int_0^{t}\|\phi_s^{\sigma_1}w(v)\|_{\sigma_2}\,ds+\int_t^{t'}\|\phi_s^{\sigma_1}w(v)\|_{\sigma_2}\,ds\\&\le\int_0^{t}\|\phi_s^{\sigma_1}v\|\,ds+\int_0^{t}\big|\,\|\phi_s^{\sigma_1}(w(v))\|_{\sigma_2}-\|(\phi_s^{\sigma_1}v)\|_{\sigma_2}\big|\,ds+\varepsilon t\max\|\cdot\|_{\sigma_2}\\&\le t(\nu_{\sigma_2}+\varepsilon K)\end{aligned} \tag{2.8}$$

where K does not depend on t and ε.

We have used the fact that $v\in A_{\varepsilon,t}$ and also that

$$D_{\sigma_1}(\phi_s^{\sigma_1}(w(v)),\phi_s^{\sigma_2}v)<\varepsilon\quad\text{for all } 0\le s\le t.$$

Let $\Gamma \in \Pi(M)$ be the free homotopy class containing γ_v. By 1.1 there exists the shortest curve with respect to σ_2 curve $\alpha \in \Gamma$ which is a closed geodesic for σ_2.

By (2.8) we have

$$l_{\sigma_2}(\alpha) \leq l_{\sigma_2}(\gamma_v) \leq t(\nu_{\sigma_2} + \varepsilon K). \tag{2.9}$$

Taking into account the estimate from below for the number of different γ_v's given by (2.7) we have

$$P^s_{\sigma_2}(t(\nu_{\sigma_2}+\varepsilon K)) \geq \frac{\varepsilon}{t(1+\varepsilon)} N_{\phi^{\sigma_1}}(t, 3\varepsilon, \tfrac{1}{2})$$

so that

$$\frac{\log P^s_{\sigma_2}(t)}{t} \geq (\nu_{\sigma_2}+\varepsilon K)^{-1} \frac{\log N_{\phi^{\sigma_1}}(t', 3\varepsilon, \frac{1}{2}) - \log t' - \log \varepsilon/(1+\varepsilon)}{t'}$$

where $t' = t(\nu_{\sigma_2}+\varepsilon K)^{-1}$. Since $t' \to \infty$ as $t \to \infty$, we have

$$P^s_{\sigma_2} \geq (\nu_{\sigma_2}+\varepsilon K)^{-1} \varliminf_{t\to\infty} (\log N_{\phi^{\sigma_1}}(t, 3\varepsilon, \tfrac{1}{2})/t),$$

and the theorem follows now from proposition 1.6. □

Theorem A follows immediately from theorem 2.1.

Theorem 2.1 and 1.5 imply the following statement.

2.2. Corollary. *Under the assumptions of* 2.1,

$$h_{\sigma_2} \geq \nu^{-1}_{\sigma_2} h^{\nu}_{\sigma_1}.$$

Let us consider now a particular case when the Liouville measure is the measure of maximal entropy for the metric σ_1. We have from theorem A and 1.4,

2.3. Corollary. *Let σ_1 be a metric of negative curvature and $h^{\lambda}_{\sigma_1} = h_{\sigma_1}$. Then for every metric σ_2 on the same manifold*

$$P^s_{\sigma_2} \geq ([\sigma_1; \sigma_2])^{-1} P_{\sigma_1} = ([\sigma_1; \sigma_2])^{-1} h_{\sigma_1}.$$

All known examples of metrics of negative curvature for which the Liouville measure has maximal entropy are of the form G/Γ where G is a symmetric space of non-compact type with real rank one and Γ is a discrete group of isometries acting freely on G. In this case the group of isometries acts transitively on the unit tangent bundle of the universal covering space G (cf. e.g. [**11**, chapter 6, theorem 6.2]). This means in particular that the Lyapunov characteristic exponents are the same for all unit tangent vectors. Let us denote the positive exponents by $\chi_1, \ldots, \chi_{n-1}$ (each exponent is counted with its multiplicity).

By Pesin's entropy formula [**23**]

$$h^{\lambda}_{\sigma} = \sum_{i=1}^{n-1} \chi_i$$

and by the above entropy estimate [**25**] and the variational principle ([**7**, § 18])

$$h_\sigma = \sup_\mu h_\mu(\phi^\sigma) \leq \sup_\mu \int \sum_{i=1}^{n-1} \chi_i(v)\, d\mu = \sum_{i=1}^{n-1} \chi_i.$$

https://doi.org/10.1017/S0143385700001656 Published online by Cambridge University Press

In the symmetric space case the entropy depends only on the universal covering space G and not on the fundamental group Γ. I am grateful to R. Spatzier who helped me to make computations.

The following list is obtained with the use of the table on pp. 532–534 of [**11**].

Symmetric space	*Dimension*	*Maximal sectional curvature*	*Entropy*
Real hyperbolic n-space (constant negative curvature)	n	$-K^2$	$(n-1)K$
Complex hyperbolic n-space	$2n$	$-K^2$	$2nK$
Quaternionic hyperbolic n-space	$4n$	$-K^2$	$(4n+2)K$
Hyperbolic plane over Cayley numbers	16	$-K^2$	$22K$

It looks like a reasonable conjecture that those are the only cases of manifolds of negative curvature for which the Liouville measure has maximal entropy.

2(B). Let us assume now that the metrics σ_1 and σ_2 are conformally equivalent, i.e. $\sigma_2=\rho\sigma_1$ where ρ is a positive scalar function. In this case $\|v\|_{\sigma_2}=(\rho(\pi_{\sigma_1}v))^{\frac{1}{2}}$ so that

$$[\sigma_1;\sigma_2]=\int_M \rho^{\frac{1}{2}}\,d\mu_{\sigma_1}. \tag{2.10}$$

Let dim $M=n$. For two conformally equivalent metrics one has

$$\mu_{\sigma_2}=\frac{v_{\sigma_1}}{v_{\sigma_2}}\rho^{n/2}\mu_{\sigma_1}$$

so that

$$\int_M \rho^{n/2}\,d\mu_{\sigma_1}=v_{\sigma_2}/v_{\sigma_1}.$$

Henceforth, by the Jensen inequality one has

$$\int_M \rho^{\frac{1}{2}}\,d\mu_{\sigma_1}\le (v_{\sigma_2}/v_{\sigma_1})^{1/n} \tag{2.11}$$

and this inequality is strict unless $\rho\equiv$ const.

Combining theorem A, 1.5, (2.10) and (2.11) we obtain

2.4. Corollary. *Let σ_1 be a Riemannian metric of negative curvature on a compact connected n-dimensional manifold M and $\sigma_2=\rho\sigma_1$. Then*

$$h_{\sigma_2}\ge P^s_{\sigma_2}\ge\left(\int_M \rho^{\frac{1}{2}}\,d\mu_{\sigma_1}\right)^{-1}h^{\lambda}_{\sigma_1}\ge (v_{\sigma_1}/v_{\sigma_2})^{1/n}h^{\lambda}_{\sigma_1}$$

and the last inequality is strict unless $\rho\equiv$ const.

2.5. Corollary. *If in addition* $h_{\sigma_1} = h^{\lambda}_{\sigma_1}$ *then among all metrics of fixed total volume conformally equivalent to the metric* σ, *the quantities* P_σ, P^s_σ *and* h_σ *reach the strict minimum at the metric proportional to* σ_1; *if we consider only* $C^{2+\delta}$ $(\delta > 0)$ *metrics of fixed volume and of negative curvature conformally equivalent to* σ_1 *then the entropy* h^{λ}_{σ} *reaches the strict maximum at the same metric.*

2(C). We will show in this subsection that the inequality between the entropies given by corollary 2.2 remains true under somewhat weaker assumptions about the 'model' metric σ_1 than the negativity of the curvature.

Let $\tilde{M}$ be the universal covering of the manifold M, $\tilde{\pi}: \tilde{M} \to M$ the covering map, $\tilde{\sigma}$ the natural lift of the Riemannian metric σ on M to $\tilde{M}$. Some of the objects associated with σ such as d_σ, π_σ, D_σ etc. can also be lifted. Naturally, we will use the symbols like $d_{\tilde{\sigma}}$, $\pi_{\tilde{\sigma}}$, $D_{\tilde{\sigma}}$ to denote those liftings.

Let for $x \in \tilde{M}$, $T > 0$, $V_{\tilde{\sigma}}(x, T)$ denote the volume of the ball of radius T about x in the metric $\tilde{\sigma}$. A. Manning [**17**] proved the following.

2.6. *For every* $x \in \tilde{M}$

$$h_\sigma \geq \lim_{T\to\infty} (\log V_{\tilde{\sigma}}(x, T)/T).$$

(*The limit exists and is independent of* x.)

2.7. Theorem. *Let* σ_1 *be a Riemannian metric on* M *such that for every* $\varepsilon > 0$ *there exists* $\kappa = \kappa(\varepsilon) > 0$ *such that if* γ_1, γ_2 *are two geodesics on* $\tilde{M}$ *of the same length* τ *and*

$$d_{\tilde{\sigma}_1}(\gamma_1(0), \gamma_2(0)) < \kappa, \quad d_{\tilde{\sigma}_1}(\gamma_1(\tau), \gamma_2(\tau)) < \kappa$$

then

$$d_{\tilde{\sigma}_1}(\gamma_1(t), \gamma_2(t)) < \varepsilon \quad \text{for every } t\colon 0 \leq t \leq \tau.$$

Then for every metric σ_2 *on* M *and every* ϕ^{σ_1}*-invariant measure* ν

$$h_{\sigma_2} \geq h^{\nu}_{\sigma_1}/\nu_{\sigma_2}.$$

Proof. The method is quite similar to the one used for the proof of theorem 2.1. First, the same argument based on (2.1) shows that it is enough to consider only ergodic measures ν. In that case we have by (2.3) for any fixed $\varepsilon > 0$ and any sufficiently large t

$$\lambda_{\sigma_1}(A_{\varepsilon,t}) > \tfrac{1}{2}, \tag{2.12}$$

where the set $A_{\varepsilon,t}$ is defined in (2.2).

Let for $\delta > 0$, $\Lambda_{t,\delta}$ be a maximal (t, δ)-separated subset of $A_{\varepsilon,t}$. Similarly to (2.6) we have

$$\operatorname{Card} \Lambda_{t,\delta} \geq N_{\phi^{\sigma_1}}(t, \delta, \lambda_{\sigma_1}(A_{\varepsilon,t})) \geq N_{\phi^{\sigma_1}}(t, \delta, \tfrac{1}{2}). \tag{2.13}$$

Let us fix a compact fundamental region $M_0 \subset \tilde{M}$ for the projection $\tilde{\pi}$ and consider for every $v \in \Lambda_{t,\delta}$ the following geodesic for the metric $\tilde{\sigma}_1$

$$\gamma_v: [-1, t+1] \to \tilde{M}$$

where $\gamma_v(s) = \pi_{\tilde{\sigma}_1} \phi^{\tilde{\sigma}_1}_s \tilde{v}$, and $\tilde{v} \in T^{\tilde{\sigma}_1}\tilde{M}$ is chosen in such a way that

$$\tilde{\pi}\tilde{v} = v \quad \text{and} \quad \gamma_v(-1) = \pi_{\tilde{\sigma}_1} \phi^{\tilde{\sigma}_1}_{-1} \tilde{v} \in M_0.$$

Since the lifting may only increase the distances on the unit tangent bundle (as well as on the manifold itself) we have for any $v_1, v_2 \in \Lambda_{t,\delta}$

$$\max_{0\le s\le t} D_{\tilde\sigma_1}(\phi_s^{\tilde\sigma_1}\tilde v_1, \phi_s^{\tilde\sigma_1}\tilde v_2) \ge \delta$$

and since any two geodesics which are C^0-close and are not very short must be C^1-close we have

$$\max_{-1\le s\le t+1} d_{\tilde\sigma_1}(\gamma_{v_1}(s), \gamma_{v_2}(s)) \ge \delta_1.$$

Therefore, by the assumption of the theorem either

$$d_{\tilde\sigma_1}(\gamma_{v_1}(-1), \gamma_{v_2}(-1)) > \kappa(\delta_1) = \delta_2 \tag{2.14}$$

or

$$d_{\tilde\sigma_2}(\gamma_{v_1}(t+1), \gamma_{v_2}(t+1)) > \delta_2. \tag{2.15}$$

Let us consider a subset $M_1 \subset M_0$ of diameter $\delta_2/2$ which contains the maximal number of points of the form $\gamma_v(-1)$ for $v \in \Lambda_{t,\delta}$. Let us denote

$$\Lambda' = \{v \in \Lambda_{t,\delta} : \gamma_v(-1) \in M_1\}.$$

For every $v_1, v_2 \in \Lambda'$ inequality (2.15) holds. On the other hand, by the maximality of M_1 there exists a positive number C_1 which depends on δ but not on t such that

$$\text{Card } \Lambda' \ge C_1 \cdot \text{Card } \Lambda_{t,\delta}. \tag{2.16}$$

Let us estimate the distance between the ends of the curve γ_v in the metric $\tilde\sigma_2$. Let us denote for the sake of brevity

$$\max_{v\in S^{\sigma_1}M} \|v\|_{\sigma_2} = \mathscr{L}^+, \quad \min_{v\in S^{\sigma_1}M} \|v\|_{\sigma_2} = \mathscr{L}^-. \tag{2.17}$$

We have for $v \in \Lambda_{t,\delta}$

$$d_{\tilde\sigma_2}(\gamma_v(-1), \gamma_v(t+1)) \le l_{\tilde\sigma_2}(\gamma_v) = \int_{-1}^{t+1} \|\phi_s^{\sigma_1}(v)\|_{\sigma_2}\, ds$$

$$\le 2\mathscr{L}^+ + \int_0^t \|\phi_s^{\sigma_1}(v)\|_{\sigma_2}\, ds \le 2\mathscr{L}^+ + t(1+\varepsilon)\nu_{\sigma_2}. \tag{2.18}$$

The last inequality follows from the definition of the set $A_{\varepsilon,t}$ (cf. (2.2)). Thus, all the points $\gamma_v(t+1)$, $v \in \Lambda'$ belong to the ball in the metric $\tilde\sigma_2$ about any point $x \in M_0$ of radius

$$t' = \text{diam}_{\tilde\sigma_2} M_0 + 2\mathscr{L}^+ + t(1+\varepsilon)\nu_{\sigma_2}. \tag{2.19}$$

By (2.15) for $v_1, v_2 \in \Lambda'$ we have

$$d_{\tilde\sigma_2}(\gamma_{v_1}(t+1), \gamma_{v_2}(t+1)) > \delta_2 \mathscr{L}^-.$$

Thus, the above mentioned ball contains Card Λ' points which lay at the fixed distance apart from each other. This means that there exists a constant $C_2 > 0$ such that

$$V_{\tilde\sigma_2}(x, t') > C_2 \cdot \text{Card } \Lambda'.$$

Furthermore, by (2.16) and (2.13)

$$V_{\tilde\sigma_2}(x, t') > C_1 C_2 \text{ Card } \Lambda_{t,\delta} \ge C_1 C_2 N_{\phi^{\sigma_2}}(t, \delta, \tfrac{1}{2}). \tag{2.20}$$

To complete the proof we use (2.19), (2.20), 2.6 and 1.6. Namely,

$$h_{\sigma_2} \geq \lim_{t' \to \infty} \frac{\log V_{\tilde{\sigma}_2}(x, t')}{t'} \geq \lim_{t \to \infty} \frac{t}{t'} \cdot \frac{\log N_{\phi^{\sigma_2}}(t, \delta, \frac{1}{2})}{t}$$
$$= \nu_{\sigma_2}^{-1}(1+\varepsilon)^{-1}(h_{\sigma_1}^{\nu} - \alpha(\delta))$$

where by 1.6 $\alpha(\delta) \to 0$ as $\delta \to 0$. Since $\varepsilon > 0$ may be chosen arbitrarily small the theorem is proved. □

The assumption of theorem 2.7 is satisfied for a metric without focal points due to the following fact.

2.8. PROPOSITION. *Let σ be a Riemannian metric on M without focal points, γ_1, γ_2: $[0, \tau] \to \tilde{M}$ be two geodesics of the same length τ on the universal covering $\tilde{M}$ parametrized by length. Then for every t: $0 \leq t \leq \tau$*

$$d_{\tilde{\sigma}}(\gamma_1(t), \gamma_2(t)) \leq d_{\tilde{\sigma}}(\gamma_1(0), \gamma_2(0)) + d_{\tilde{\sigma}}(\gamma_1(\tau), \gamma_2(\tau)).$$

Proof. For any metric without focal points the distance between corresponding points on the geodesic rays issuing from any given point is increasing. Let γ' be the geodesic connecting the points $\gamma_1(0)$ and $\gamma_2(\tau)$ and let $l_{\tilde{\sigma}}(\gamma') = \rho$. By the above-mentioned property and the triangle inequality one has (cf. figure 1).

$$d_{\tilde{\sigma}}(\gamma_1(t), \gamma_2(t)) \leq d_{\tilde{\sigma}}\left(\gamma_1(t), \gamma'\left(\frac{ts}{\tau}\right)\right) + d_{\tilde{\sigma}}\left(\gamma'\left(\frac{ts}{\tau}\right), \gamma_2(t)\right)$$
$$\leq d_{\tilde{\sigma}}(\gamma_1(\tau), \gamma_2(\tau)) + d_{\tilde{\sigma}}(\gamma_1(0), \gamma_2(0)).$$ □

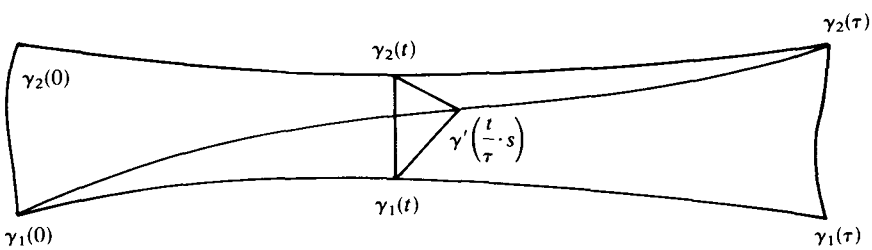

FIGURE 1

Theorem 2.7 and proposition 2.8 immediately imply the following

2.9. COROLLARY. *If σ_1 is a metric without focal points then for every metric σ_2 on the same manifold and every ϕ^{σ_1}-invariant measure ν*

$$h_{\sigma_2} \geq h_\nu(\phi^{\sigma_1})/\nu_{\sigma_2}.$$

In particular $h_{\sigma_2} \geq [\sigma_1; \sigma_2]^{-1} h_{\sigma_1}^{\lambda}$.

2.10. *Remark.* It follows easily from proposition 2.8 that for every metric σ without focal points

$$h_\sigma = \lim_{T \to \infty} (V_{\tilde{\sigma}}(x, T)/T)$$

for every $x \in \tilde{M}$.

2(D). The proof of theorem 2.7 involves an estimate of $\tilde{\sigma}_2$-distance between two ends of a 'typical' long $\tilde{\sigma}_1$-geodesic on $\tilde{M}$, (cf. (2.18)). This estimate which depends only on ergodicity of the geodesic flow (here and later in this section we will simply refer to 'ergodicity' meaning the ergodicity with respect to Liouville measure) is interesting in itself; it can be reformulated in terms of an individual (but still typical) geodesic.

Namely, let for $v \in S^{\sigma_1}M$, $t > 0$

$$d_{\sigma_1,\sigma_2}(v, t) = d_{\tilde{\sigma}_2}(\tilde{\pi}\tilde{v}, \tilde{\pi}(\phi_t^{\tilde{\sigma}_1}\tilde{v}))$$

where $\tilde{v}$ is an arbitrary lift of v into $S^{\tilde{\sigma}_1}\tilde{M}$.

Obviously,

$$d_{\sigma_1,\sigma_2}(v, t) \le l_{\sigma_1,\sigma_2}(v, t) \overset{\text{def}}{=} \int_0^t \|\phi_s^{\sigma_1}(v)\|_{\sigma_2}\, ds \tag{2.21}$$

so that we have from the Birkhoff ergodic theorem:

2.11. *If the geodesic flow ϕ^{σ_1} is ergodic, then for almost every $v \in S^{\sigma_1}M$*

$$\overline{\lim_{t\to\infty}}\,(d_{\sigma_1,\sigma_2}(v, t)/t) \le [\sigma_1; \sigma_2].$$

(As in the case of ergodicity 'almost every' means 'except for a set of measure zero with respect to Liouville measure'; this notion, indeed, is determined by the smooth structure on $S^{\sigma}M$.) This inequality can be refined in several ways which in turn allows us to obtain stronger versions of the statement of theorem 2.7.

Let $\Sigma \subset S^{\sigma_1}M$ be a transversal section for the geodesic flow ϕ^{σ_1}, $\phi_{\sigma_1}^{\Sigma}$ the Poincare map induced by the geodesic flow on Σ, $\lambda_{\sigma_1}^{\Sigma}$ the $\phi_{\sigma_1}^{\Sigma}$-invariant measure on Σ induced by the Liouville measure λ_{σ_1} and for $w \in \Sigma$ let $t_\Sigma(w)$ be the return time so that

$$\phi_{\sigma_1}^{\Sigma}(w) = \phi_{t_\Sigma(w)}^{\sigma_1}(w).$$

Furthermore, let

$$d_{\sigma_1,\sigma_2}^{\Sigma}(w) = d_{\tilde{\sigma}_2}(\tilde{\pi}\tilde{w}, \tilde{\pi}\phi_{t_\Sigma(w)}^{\tilde{\sigma}_1}\tilde{w})$$

be the $\tilde{\sigma}_2$ distance between two successive intersections of the $\tilde{\sigma}_1$ geodesic with the lifts of Σ (here, as before $\tilde{w}$ is an arbitrary lifting of w) and

$$l_{\sigma_1,\sigma_2}^{\Sigma}(w) = \int_0^{t_\Sigma(w)} \|\phi_s^{\sigma_1}(w)\|_{\sigma_2}\, ds,$$

be the $\tilde{\sigma}_2$ length of the $\tilde{\sigma}_1$-geodesic connecting those intersections. Clearly

$$d_{\sigma_1,\sigma_2}^{\Sigma}(w) \le l_{\sigma_1,\sigma_2}^{\Sigma}(w) \tag{2.22}$$

and the equality in (2.22) means that σ_1-geodesic $\{\phi_s^{\sigma_1}w\}_{s=0}^{t_\Sigma(w)}$ is also a σ_2-geodesic.

Since we assume that the geodesic flow ϕ^{σ_1} is ergodic, almost every σ_1-geodesic intersects the section Σ infinitely many times; we will consider only such geodesics.

The distance $d_{\sigma_1,\sigma_2}(v, t)$ does not exceed the sum of the distances between successive intersections of the orbit $\{\phi_s^{\tilde{\sigma}_1}v\}$ with the lift of Σ, from the last intersection for negative s to the first one for $s \ge t$. For almost every v the asymptotic behaviour of this sum is given by the Birkhoff ergodic theorem. Namely, we have

2.12. *If ϕ^{σ_1} is ergodic then for almost every v*

$$\overline{\lim_{t\to\infty}} \frac{d_{\sigma_1,\sigma_2}(v,t)}{l_{\sigma_1,\sigma_2}(v,t)} \le \frac{\int_\Sigma d^\Sigma_{\sigma_1,\sigma_2}(w)\,d\lambda^\Sigma_{\sigma_1}(w)}{\int_\Sigma l^\Sigma_{\sigma_1,\sigma_2}(w)\,d\lambda^\Sigma_{\sigma_1}(w)} \overset{\text{def}}{=} D_\Sigma$$

where by (2.22) $D_\Sigma \le 1$.

Taking into account the definition of $l_{\sigma_1,\sigma_2}(v,t)$ (cf. (2.21)) and using the Birkhoff theorem once more, we have

2.13. *For almost every v*

$$\overline{\lim_{t\to\infty}} \frac{d_{\sigma_1,\sigma_2}(v,t)}{t} \le D_\Sigma \cdot [\sigma_1;\sigma_2].$$

In the proof of theorem 2.7 instead of the set $A_{\varepsilon,t}$ one can start from the set of $v \in S^{\sigma_1}M$ for which

$$d_{\sigma_1,\sigma_2}(v,t) < t(D_\Sigma \cdot [\sigma_1;\sigma_2] + \varepsilon)$$

and proceed in the same way. This leads to the following refinement of theorem 2.7.

2.14. *Under the assumptions of theorem* 2.7

$$h_{\sigma_2} \ge (D_\Sigma[\sigma_1;\sigma_2])^{-1} h^\lambda_{\sigma_2}$$

where Σ is an arbitrary section of the geodesic flow ϕ^{σ_2}.

2.15. PROPOSITION. *If there is at least one σ_1-geodesic which is not a σ_2-geodesic then for almost every v*

$$\overline{\lim_{t\to\infty}} (d_{\sigma_1,\sigma_2}(v,t)/t) < [\sigma_1;\sigma_2].$$

(Here geodesics are considered as curves on M without parametrization.)

Proof. Let $\{\phi^{\sigma_1}_t v\}_{t=0}^{t_0}$ be a sufficiently short orbit of ϕ^{σ_1} whose projection on M is not a σ_2-geodesic. Let Σ be a section for ϕ^{σ_1} consisting of two small hypersurfaces in $S^{\sigma_1}M$ near the vectors v and $\phi^{\sigma_1}_{t_0}v$. Then

$$d^\Sigma_{\sigma_1,\sigma_2}(v) < l^\Sigma_{\sigma_2,\sigma_2}(v) \tag{2.23}$$

and by the continuity of the functions $d^\Sigma_{\sigma_1,\sigma_2}$ and $l^\Sigma_{\sigma_1,\sigma_2}$ at v this inequality holds in a neighbourhood of v. This together with (2.22) imply that $D_\Sigma < 1$. □

We obtain immediately from 2.14 and 2.15

2.16. COROLLARY. *If the assumptions of* 2.7 *and* 2.15 *are satisfied then*

$$h_{\sigma_2} > ([\sigma_1;\sigma_2])^{-1} h^\lambda_{\sigma_1}.$$

2.17. *Remark.* It is not difficult to show that for conformally equivalent metrics the assumption of 2.15 is always satisfied unless the metrics are proportional.

2.18. PROPOSITION. *Let us assume that σ_2 is a metric of negative curvature. Let Σ_n be an arbitrary sequence of sections for ϕ^{σ_1} such that $\min_{v\in\Sigma_n} t_{\Sigma_n}(v)\to\infty$ as $n\to\infty$.*

Then, for almost every $v\in S^{\sigma_1}M$

$$\lim_{t\to\infty}\frac{d_{\sigma_1,\sigma_2}(v,t)}{l_{\sigma_1,\sigma_2}(v,t)}=\lim_{n\to\infty} D_{\Sigma_n}.$$

Proof. The $\tilde\sigma_2$ geodesic connecting any two points on $\tilde M$ is unique and by Morse's lemma [**22**] stays within a bounded distance $r=r(\sigma_1,\sigma_2)$ from the $\tilde\sigma_1$-geodesic connecting the same points. Let us fix $v\in S^{\sigma_1}M$ and $t>0$, take any lift $\tilde v$ of v, denote by γ_1 the $\tilde\sigma_1$-geodesic of length t starting at $\tilde v$ and by γ_2 the $\tilde\sigma_2$-geodesic connecting the ends of γ_1 (see figure 2).

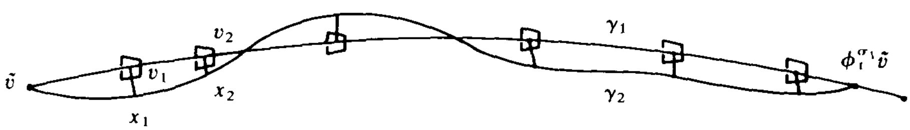

FIGURE 2

Let Σ be a section for the geodesic flow ϕ^{σ_1} and $v_1,\ldots,v_{k(t)}$ be the successive points of intersection (in $S^{\tilde\sigma_1}\tilde M$) of γ_1 with the lifts of Σ. Furthermore, let x_i be the orthogonal projection of the point $\tilde\pi v_i$ to γ_2. We mean here projection with respect to $\tilde\sigma_2$; such a projection exists and is unique because σ_2 is a metric of negative curvature. We have by the triangle inequality

$$d_{\tilde\sigma_2}(x_i,x_{i+1})\geq d_{\tilde\sigma_2}(\tilde\pi v_i,\tilde\pi v_{i+1})-2r=d^{\Sigma}_{\sigma_1,\sigma_2}-2r.$$

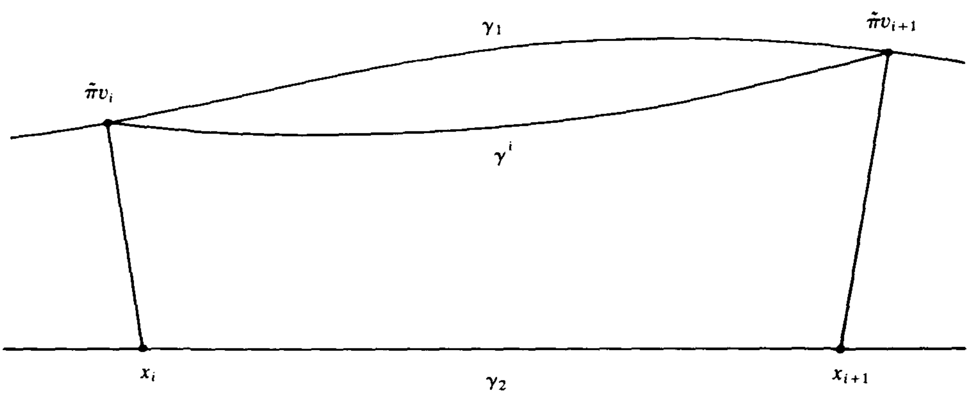

FIGURE 3

Let γ^i be the $\tilde\sigma_2$ geodesic connecting the points $\tilde\pi v_i$ and $\tilde\pi v_{i+1}$ (cf. figure 3). We have (see (2.17)):

$$\begin{aligned}t_\Sigma(v_i)=d_{\tilde\sigma_1}(\tilde\pi v_i,\tilde\pi v_{i+1})&\leq l_{\tilde\sigma_1}(\gamma^i)\leq(\mathscr{L}^-)^{-1}l_{\tilde\sigma_2}(\gamma^i)\\&=(\mathscr{L}^-)^{-1}d_{\tilde\sigma_2}(\tilde\pi v_i,\tilde\pi v_{i+1})\\&=(\mathscr{L}^-)^{-1}d^{\Sigma}_{\sigma_1,\sigma_2}(v_i)\end{aligned}$$

so that

$$d^{\Sigma}_{\sigma_1,\sigma_2}(v_i) \geq \min t_\Sigma \cdot \mathscr{L}^-.$$

Thus, we have for $i = 1, \ldots, k(t)-1$

$$\frac{d_{\tilde{\sigma}_2}(x_i, x_{i+1})}{d_{\tilde{\sigma}_2}(\tilde{\pi}_1 v_i, \tilde{\pi}_1 v_{i+1})} > \frac{\mathscr{L}^- \min t_\Sigma - 2r}{\mathscr{L}^- \min t_\Sigma}.$$

If $\min t_\Sigma$ is big enough this implies in particular that the points $x_1, \ldots, x_{k(t)}$ lie on γ_2 in the same order as $\tilde{\pi}v_1, \ldots, \tilde{\pi}v_{k(t)}$ on γ_1, so that

$$\frac{d_{\tilde{\sigma}_2}(x_1, x_{k(t)})}{\sum_{i=1}^{k(t)-1} d_{\tilde{\sigma}_2}(\tilde{\pi}v_i, \tilde{\pi}v_{i+1})} = \frac{\sum_{i=1}^{k(t)-1} d_{\tilde{\sigma}_2}(x_i, x_{i+1})}{\sum_{i=0}^{k(t)-2} d^{\Sigma}_{\sigma_1,\sigma_2}(\phi^{\Sigma}_{\sigma_1} v_1)} > \frac{\mathscr{L}^- \cdot \min t_\Sigma - 2r}{\mathscr{L}^- \cdot \min t_\Sigma}. \tag{2.24}$$

Now we can proceed to the ratio in question. We have

$$\begin{aligned}\frac{d_{\sigma_1,\sigma_2}(v,t)}{l_{\sigma_1,\sigma_2}(v,t)} &= \frac{d_{\tilde{\sigma}_2}(x_1, x_{k(t)}) + d_{\tilde{\sigma}_2}(\tilde{\pi}\tilde{v}, x_1) + d_{\tilde{\sigma}_2}(x_{k(t)}, \pi\phi^{\tilde{\sigma}_1}_t \tilde{v})}{\sum_{i=1}^{k(t)-1} d_{\tilde{\sigma}_2}(\tilde{\pi}v_i, \tilde{\pi}v_{i+1})} \\ &\quad \times \frac{\sum_{i=0}^{k(t)-2} d^{\Sigma}_{\sigma_1,\sigma_2}((\phi^{\Sigma}_{\sigma_1})^i v_1)}{\sum_{i=0}^{k(t)-2} l^{\Sigma}_{\sigma_1,\sigma_2}((\phi^{\Sigma}_{\sigma_1})^i v_1)} \times \frac{l_{\tilde{\sigma}_2}(\gamma'_1)}{l_{\tilde{\sigma}_2}(\gamma_1)}\end{aligned} \tag{2.25}$$

where γ'_1 is a segment of γ_1 between v_1 and $v_{k(t)}$. By the Birkhoff ergodic theorem for almost every $v \in S^{\sigma_1}M$ as $t \to \infty$ the time before the first intersection of γ_1 with Σ and after the last such intersection becomes negligible in comparison with t. This easily implies that the third multiple in (2.25) tends to 1 for such v and also that the first multiple becomes close to the expression at the right-hand side of (2.24). The second multiple goes to D_Σ for almost every v.

Thus, we obtain from 2.12 and (2.25)

$$D_\Sigma \cdot \frac{\mathscr{L}^- \min t_\Sigma - 2r}{\mathscr{L}^- \min t_\Sigma} \leq \varliminf_{t\to\infty} \frac{d_{\sigma_1,\sigma_2}(v,t)}{l_{\sigma_1,\sigma_2}(v,t)} \leq \varlimsup_{t\to\infty} \frac{d_{\sigma_1,\sigma_2}(v,t)}{l_{\sigma_1,\sigma_2}(v,t)} \leq \inf_\Sigma D_\Sigma. \tag{2.26}$$

Since $\mathscr{L}^-$ and r do not depend on Σ, v, t the proposition follows from (2.26). □

3. *Two-dimensional case*

3(A). Let now M be a compact surface with negative Euler characteristic. By 1.7 we can associate with any given Riemannian metric σ a metric of constant negative curvature σ_0 and the number

$$\rho_\sigma \overset{\text{def}}{=} [\sigma; \sigma_0] = \int_M \rho^{\frac{1}{2}}\, d\mu_{\sigma_0}.$$

We will show soon that this number may serve as a measure of deviation of σ from the metrics of constant negative curvature.

Let us also remark that in the two-dimensional case, for any two conformally equivalent metrics σ_1 and $\sigma_2 = \rho\sigma_1$ of the same total area one has

$$\begin{aligned}[\sigma_1; \sigma_2] &= \int_M \rho^{\frac{1}{2}}\, d\mu_{\sigma_1} = \int_M \rho^{-\frac{1}{2}} \rho\, d\mu_{\sigma_1} = \int_M (\rho^{-1})^{\frac{1}{2}}\, d\mu_{\sigma_2} \\ &= [\sigma_2; \sigma_1].\end{aligned} \tag{3.1}$$

https://doi.org/10.1017/S0143385700001656 Published online by Cambridge University Press

The Gauss–Bonnet formula says that the curvature $-K^2$ for a metric of constant negative curvature is determined by the Euler characteristic E and the total area v, namely

$$-K^2 = -K^2(E, v) = 2\pi E/v.$$

Henceforth by 1.2, 1.4 and the computation of the entropy in constant negative curvature case, we have for every such metric σ_0

$$P^s_{\sigma_0} = P_{\sigma_0} = h_{\sigma_0} = h^\lambda_{\sigma_0} = K(E, v) = \left(\frac{-2\pi E}{v}\right)^{\frac{1}{2}}. \tag{3.2}$$

The following theorem is an immediate consequence of corollary 2.3, (3.1), (3.2) and 1.8.

3.1. Theorem. *For every Riemannian metric σ on a compact surface of negative Euler characteristic E*

$$P_\sigma \geq h_\sigma \geq P^s_\sigma \geq \rho_\sigma^{-1}\left(\frac{-2\pi E}{v_\sigma}\right)^{\frac{1}{2}} \geq \left(\frac{-2\pi E}{v_\sigma}\right)^{\frac{1}{2}}$$

and the last inequality is strict unless σ is a metric of constant negative curvature.

3.2. Theorem. *For a metric σ without focal points*

$$h^\lambda_\sigma \leq \rho_\sigma\left(\frac{-2\pi E}{v_\sigma}\right)^{\frac{1}{2}} \leq \left(\frac{-2\pi E}{v_\sigma}\right)^{\frac{1}{2}}$$

and both inequalities are strict for every metric with non-constant curvature.

Proof. By 1.7 we can find a metric σ_0 of constant negative curvature such that $\sigma_0 = \rho^{-1}\sigma$. By theorem 2.7 and (3.1)

$$\left(\frac{-2\pi E}{v_{\sigma_0}}\right)^{\frac{1}{2}} = h_{\sigma_0} \geq ([\sigma;\sigma_0])^{-1} h^\lambda_\sigma = ([\sigma_0;\sigma)]^{-1} h^\lambda_\sigma = \rho_\sigma^{-1} h^\lambda_\sigma.$$

If σ is not a metric of constant negative curvature then ρ is not a constant. Consequently $\rho_\sigma < 1$ and by corollary 2.16 and remark 2.17, $h^\lambda_\sigma > \rho_\sigma h_{\sigma_0}$. □

Theorem B follows immediately from theorems 3.1 and 3.2.

3.3. Corollary. *If for a metric σ without focal points on a surface with negative Euler characteristic*

$$h_\sigma = h^\lambda_\sigma$$

then σ is a metric of constant negative curvature.

There is a good chance that the assertion of theorem 3.2 is true for a wider class of Riemannian metrics without conjugate points. On the other hand, it is at least highly probable that for general Riemannian metrics of fixed total area the entropy with respect to Liouville measure may be arbitrarily large (although at the present time no example is known of a metric with conjugate points and positive h^λ_σ). Thus, the metrics of constant negative curvature occupy a minimax position among all metric without focal points of fixed total volume v; the number $K(E, v)$, which is the common value of the topological entropy and the entropy with respect to the Liouville measure for metrics of constant negative curvature, separates the values of the two entropies for all other metrics.

3(B). Theorems 3.1 and 3.2 also support the idea that ρ_σ is a good measure of deviation of σ from the metrics of constant negative curvature. It is an interesting problem to estimate this number through different geometric characteristics of the metric.

As an example of such a result which makes use of theorem 3.2 we present an estimate of ρ_σ through the volume and injectivity radius

$$R_\sigma = \min_{x\in M} \sup\{r\colon \exp_x \text{ is regular and invertible on } r\text{-ball around the origin in } T_xM\}.$$

In particular, if $d_\sigma(x, y) < R_\sigma$ then the shortest geodesic connecting x and y is unique.

3.4. THEOREM. *For any Riemannian metric σ on a compact surface with negative Euler characteristic*

$$\rho_\sigma \geq \sup_{0<\alpha<1} \left(\frac{-\pi E}{2}\right)^{\frac{1}{2}} \cdot \alpha \cdot \frac{R_\sigma}{v_\sigma^{1/2}} \cdot \left(\log \frac{v_\sigma^{1/2}}{R_\sigma} - \log \frac{1-\alpha}{4}\right)^{-1}.$$

If, moreover, the curvature of σ is greater than or equal to Δ then

$$\rho_\sigma \geq \sup_{0<\alpha<1} (-\pi E/2)^{\frac{1}{2}} \alpha R_\sigma (\log(\pi \sin h \tfrac{1}{2}|\Delta|^{\frac{1}{2}}) - \log(1-\alpha)|\Delta|^{\frac{1}{2}}/8)^{-1}.$$

Let for a given metric σ, $x \in M$ and $r > 0$, $B_r(x)$ be the ball of radius r about x, $V_r(x)$ be the volume of this ball and if $r < R_\sigma$ let $L_r(x)$ be the circle of radius r around x i.e.

$$L_r(x) = \{y \in M\colon d_\sigma(y, x) = r\} \quad \text{and} \quad l_r(x) = l_\sigma(L_r(x)).$$

3.5. LEMMA. (cf. [3]). *If M is a compact surface different from the sphere then for any $x \in M$, $r \leq R_\sigma$*

$$V_r(x) > r^2.$$

Proof. Since for $r \leq R_\sigma$

$$V_r(x) = \int_0^r l_s(x)\, ds$$

it is enough to prove that for $r < R_\sigma$

$$l_r(x) > 2r. \tag{3.3}$$

Let $y \in L_r(x)$. We will prove that

$$\max_{z\in L_r(x)} d_\sigma(y, z) > r. \tag{3.4}$$

Inequality (3.3) follows from (3.4) because $L_r(x) = \partial B_r(x)$ is diffeomorphic to a circle and this circle consists of two arcs connecting points y, z such that $d_\sigma(y, z) > r$. Consequently the length of each arc is greater than r.

Suppose that (3.4) is false, i.e.

$$L_r(x) \subset B_r(y). \tag{3.5}$$

On the other hand, $x \in \partial B_r(y) = L_r(y)$ so that arbitrarily close to x one can find a point $w \in M \backslash B_r(y)$. Since $B_r(y)$ is diffeomorphic to a disk the complement $M\backslash B_r(y)$ is connected and by (3.5)

$$M\backslash B_r(y) \subset B_r(x). \tag{3.6}$$

Since $\partial(M\backslash B_r(y)) = \partial B_r(y) = L_r(y)$ is a topological circle then by (3.6) $M\backslash B_r(y)$ is homeomorphic to a disk and M is homeomorphic to a sphere being a union of two disks, $B_r(x)$ and $M\backslash B_r(y)$, with common boundary. □

Remark. For the sphere the statement of the lemma is still true if $r \le R_\sigma/2$. In this case inclusion (3.6) is impossible because $B_r(x) \subset B_{2r}(y) \subset B_{R_\sigma}(y)$ and $M\backslash B_{R_\sigma}(y) \neq \varnothing$.

3.6. PROPOSITION. *Let* $0<\alpha<1$. *Then*

$$P_\sigma^s < \frac{\log (v_\sigma/R_\sigma^2) - 2\log((1-\alpha)/4)}{\alpha R_\sigma}.$$

If the curvature of σ *is greater than* Δ *then*

$$P_\sigma^s < \frac{2((\log(\pi \sin h \frac{1}{2}|\Delta|^{\frac{1}{2}}) - \log((1-\alpha)/8)|\Delta|^{\frac{1}{2}}))}{\alpha R_\sigma}.$$

Proof. Let us fix a number c, $0<c<1$ and construct a maximal set $B_c \subset M$ such that the distance between any two of its points is greater than or equal to cR_σ. The balls of radius $cR_\sigma/2$ around the points of B_c are disjoint so that by lemma 3.5

$$\text{Card } B_c < v_\sigma/((c/2)R_\sigma)^2. \tag{3.7}$$

On the other hand, the maximality of B_c implies that the balls of radius cR_σ around the points of B_c cover M.

Let now $\gamma \in \mathcal{P}_\sigma(T)$. Let us fix $\alpha: 0<\alpha<1$ and divide the geodesic γ by points $p_0, p_1, \ldots, p_{m-1}, p_m = p_0$ into the segments of length αR_σ (the last segment $(p_{m-1}p_0)$ may be shorter).

Let for $i = 0, 1, \ldots, m-1$, $q_i \in B_c$ be such a point that $d_\sigma(p_i, q_i) < cR_\sigma$.

If $\alpha + 2c \le 1$ then all four points p_i, p_{i+1}, q_i, q_{i+1} as well as the shortest geodesics connecting these points lie in a ball of radius R_σ.

In particular, the geodesic segment (p_ip_{i+1}) will be homotoped to the segment (q_iq_{i+1}) in such a manner that the ends move uniformly along the geodesic segments (p_iq_i) and $(p_{i+1}q_{i+1})$ respectively. This means that the curve γ is homotopical to a piecewise geodesic curve $\gamma' = (q_0, q_1, q_2, \ldots, q_{m-1}, q_0)$ where $q_i \in B_c$, $i = 0, \ldots, m-1$. Obviously, if γ_1, γ_2 belong to different elements of Π the corresponding curves γ_1' and γ_2' must be different.

Since $m = [l_\sigma(\gamma)/\alpha R_\sigma] + 1 \le [T/\alpha R_\sigma] + 1$ we have

$$P_\sigma^s(T) \le (\text{Card } B_c)^{(T/\alpha R_\sigma)+1} \tag{3.8}$$

and by (3.7) and (3.8)

$$P_\sigma^s < \frac{1}{\alpha R_\sigma} \log (v_\sigma/((c/2)R_\sigma)^2). \tag{3.9}$$

Putting in (3.9) $c = (1-\alpha)/2$ we obtain the first statement of the proposition.

In order to obtain the second statement we remark that in our approximation procedure the pair (q_i, q_{i+1}) cannot be an arbitrary pair of elements of the set B_c because $d_\sigma(q_{i+1}, q_i) < R_\sigma$. Let K_c be the maximum number of elements of B_σ in a ball of radius R_σ. Instead of (3.8) we have then

$$P_\sigma^s(T) < (K_c)^{[T/\alpha R_\sigma]+1}. \tag{3.10}$$

Clearly

$$K_c < \frac{\text{maximal volume of an } R_\sigma\text{-ball}}{\text{minimal volume of a } cR_\sigma/2 \text{ ball}}.$$

By lemma 3.5 and the above volume estimate for metrics with the curvature bounded from below ([**5**, chapter 11, theorem 15], we have

$$K_c < \pi\left(\frac{\sinh\left(\frac{1}{2}|\Delta|^{\frac{1}{2}}\right)}{\frac{c}{4}|\Delta|^{\frac{1}{2}}}\right)^2. \tag{3.11}$$

Now the second statement of 3.6 follows immediately from (3.10) and (3.11). □

Proof of 3.4. By theorem 3.1 and proposition 3.6 we have for any $\alpha: 0<\alpha<1$

$$\rho_\sigma \geq \left(\frac{-2\pi E}{v_\sigma}\right)^{\frac{1}{2}}(P^s_\sigma)^{-1} > \left(\frac{-\pi E}{v_\sigma}\right)^{\frac{1}{2}} \cdot \alpha\,\frac{R_\sigma}{v_\sigma^{1/2}} \cdot \left(\log\frac{v_\sigma^{1/2}}{R_\sigma} - \log\frac{1-\alpha}{4}\right)^{-1}.$$

With the curvature assumption we have similarly,

$$\rho_\sigma > \left(\frac{-\pi E}{v_\sigma}\right)^{\frac{1}{2}} \alpha R_\sigma \left(\log\left(\pi \sin h\tfrac{1}{2}|\Delta|^{\frac{1}{2}}\right) - \log\frac{1-\alpha}{8}|\Delta|^{\frac{1}{2}}\right)^{-1}. \qquad □$$

Remark. I am almost certain that there exists an estimate for P^s_σ from above depending only on R_σ without the curvature assumption but I have not been able to find it.

I would like to mention two more results related to the discussed topic which were proved by D. Epstein and A Douady correspondingly in response to my questions.

3.7. *There exists a constant $D = D(E, v, R)$ such that if σ_0 is a metric of constant negative curvature with total volume v and injectivity radius R on a surface with negative Euler characteristic E and $\sigma = \rho\sigma_0$ is a metric of non-positive curvature then*

$$\rho_\sigma > D(E, v, R)$$

(D. Epstein).

3.8. *Given $\varepsilon > 0$, $E < 0$ there exists a metric σ_0 of constant negative curvature and total volume 1 on a surface with Euler characteristic E and a positive function ρ with $\int \rho\, d\mu_{\sigma_0} = 1$ such that $\sigma = \rho\sigma_0$ is a metric of negative curvature and $\rho_\sigma < \varepsilon$* (A. Douady).

3(C). Another natural measure of deviation from the constant curvature, at least for metrics of non-positive curvature, seems to be the average of the absolute value of the square root of curvature K_σ.

By the Gauss–Bonnet theorem

$$\int_M K_\sigma\, d\mu_\sigma = 2\pi E/v_\sigma$$

https://doi.org/10.1017/S0143385700001656 Published online by Cambridge University Press

so that if $K_\sigma \le 0$ then by the Jensen inequality

$$k_\sigma \overset{\text{def}}{=} \int_M (-K_\sigma)^{\frac{1}{2}}\, d\mu_\sigma \le (-2\pi E/v_\sigma)^{\frac{1}{2}} = k_{\sigma_0} \tag{3.12}$$

where σ_0 is any metric of constant negative curvature on M with the same total area as σ.

Recently A. Manning [**18**] and independently P. Sarnak [**26**] proved the following entropy estimate from below.

3.9. *If σ is a metric of negative curvature then*

$$h_\sigma^\lambda \ge k_\sigma.$$

(Sarnak also has a multidimensional version of this inequality.)

Combining this result with our estimate for the metric entropy from above (theorem 3.2) one obtains an inequality between the two measures of deviation from constant curvature. Let, as above, σ_0 be any metric of constant negative curvature such that $v_{\sigma_0} = v_\sigma$.

3.10. COROLLARY. *For any metric σ of negative curvature*

$$\rho_\sigma \le k_\sigma(-2\pi E/v_\sigma)^{-\frac{1}{2}} = k_\sigma / k_{\sigma_0}$$

and this inequality is strict for every metric of non-constant curvature.

Some examples show that the gap between the two sides of the last inequality may be quite big. Namely, one can construct on every surface with negative Euler characteristic a metric σ of negative curvature and of total area 1 such that k_σ is arbitrarily small but the entropy h_σ^λ (and consequently ρ_σ) is bounded from below by a fixed positive number.

4. *Counting hyperbolic geodesics*

4(A). In our proofs of the results dealing with closed geodesics (theorem 2.1 and its corollaries 2.2, 2.3 and theorem 3.1) we used ergodic properties of geodesic flows only for manifolds of negative curvature. In other words, we used ergodic theory for Anosov flows. Closed geodesics for a general metric were produced by a variational argument. In the two-dimensional case, since these geodesics are the shortest in their free homotopy classes, they may be either hyperbolic or parabolic (degenerate) [**9**, § 4.6]. Variational methods do not allow one to distinguish between these two cases.

Now we are going to apply ergodic theory for general smooth dynamical systems in order to produce an estimate from below for the growth rate of hyperbolic closed geodesics and more specifically, the geodesics with a fixed degree of hyperbolicity.

For low-dimensional dynamical systems (dimension two for discrete time and three for continuous time) positive entropy implies the existence of large sets of orbits with hyperbolic behaviour, including many periodic hyperbolic orbits [**13**],

[16]. We refer to [13] for all background information, including the discussion on the Lyapunov characteristic exponents. Let us note that the Lyapunov exponents for a flow coincide with the corresponding exponents for its time-one map; a trivial zero exponent, corresponding to the flow direction should not be counted.

We will consider the following situation which includes as a special case the geodesic flow for a Riemannian metric of class $C^{2+\delta}$ ($\delta > 0$) on a compact surface.

Let N be a compact 3-dimensional manifold, $f = \{f_t\}_{t\in\mathbf{R}}$ a $C^{1+\delta}$ ($\delta > 0$) flow on N without fixed points, μ an f-invariant Borel probability ergodic measure with non-zero Lyapunov exponents $\chi_1^\mu < 0 < \chi_2^\mu$ (cf. [13]). Let for α, $\beta > 0$, $P_{\alpha,\beta}(f, T)$ be the number of hyperbolic periodic orbits of period $\leq T$ with a negative characteristic exponent $\leq -\alpha$ and a positive exponent $\geq \beta$.

4.1. THEOREM. *For every* $\varepsilon > 0$

$$\underline{\lim}_{T\to\infty} \log (P_{-\chi_1^\mu+\varepsilon,\chi_2^\mu-\varepsilon}(f, T))/T \geq h_\mu(f).$$

This theorem is essentially a continuous-time version of the two-dimensional case of theorem 4.3 from [10]. (The multi-dimensional version of theorem 4.1 is also true, but we do not need it for our purpose.) The only difference is that the degree of hyperbolicity of produced periodic orbits is not estimated in [13] explicitly. The changes in the proof which provide such an estimate are very transparent. As for the translation from diffeomorphisms to flows, it can be done straightforwardly beginning from the definition of the Lyapunov metric through the flow versions of propositions 2.3 and 2.4 which characterize the behaviour of orbits near a regular orbit to main closing type lemma and the final arguments which use our proposition 1.6 instead of its discrete-time version. As a matter of fact, the result for diffeomorphisms can be deduced from the result for flows by means of the standard suspension construction.

Returning to the case of Riemannian metric σ on a compact surface, let us denote for $\chi > 0$, $T > 0$.

$\mathcal{P}_{\sigma,\chi}(T)$ – the set of $\gamma \in \mathcal{P}_\sigma(T)$ such that γ is hyperbolic and its positive Lyapunov exponent is $\geq \chi$ (then the negative exponent is $\leq -\chi$, because the two exponents have the same absolute value);

$$P_{\sigma,\chi}(T) = \operatorname{Card} \mathcal{P}_{\sigma,\chi}(T);$$

$$P_{\sigma,\chi} = \underline{\lim}_{T\to\infty} (\log (\min (P_{\sigma,\chi}(T), 1)))/T. \tag{4.1}$$

4.2. COROLLARY. *For every Riemannian metric* σ *of class* $C^{2+\delta}$ ($\delta > 0$) *on a compact surface and any* $\varepsilon > 0$

$$P_{\sigma,h_\sigma-\varepsilon} \geq h_\sigma.$$

The proof goes exactly the same way as the proof of corollary (4.4) in [13]. Namely, assuming that $h_\sigma > 0$ one can find an ergodic ϕ^σ-invariant measure μ_ε such that

$$h_{\mu_\varepsilon}(\phi^\sigma) > h_\sigma - \varepsilon.$$

On the other hand, by the above entropy estimate [**25**] we have for the Lyapunov exponents $\chi_1^{\mu_\varepsilon}$, $\chi_2^{\mu_\varepsilon} = -\chi_1^{\mu_\varepsilon} > 0$ of the geodesic flow with respect to μ_ε

$$\chi_2^{\mu_\varepsilon} \geq h_{\mu_\varepsilon}(\phi^\sigma).$$

Now theorem 4.1 applied to the geodesic flow provides for every $\varepsilon_1 < \varepsilon$

$$P_{\sigma, h_\sigma - \varepsilon} \geq P_{\sigma, h_\sigma - \varepsilon_1} \geq h_{\mu_{\varepsilon_1}}(\phi^\sigma) > h_\sigma - \varepsilon_1. \qquad \square$$

Theorem 3.1 and corollary 4.2 imply the absolute estimate for the growth rate of hyperbolic closed geodesics.

4.3. COROLLARY. *For any Riemannian metric σ of class $C^{2+\delta}$ $(\delta > 0)$ on a compact surface with negative Euler characteristic E*

$$P_{\sigma, K(E, v_\sigma)} \geq K(E, v_\sigma) = (-2\pi E / v_\sigma)^{\frac{1}{2}}$$

and this inequality is strict unless σ is a metric of constant negative curvature.

Since for metrics of constant negative curvature $-K^2$, the positive exponent of every geodesic is equal to K (and, consequently, the negative one is equal to $-K$) we can say that metrics of non-constant curvature have more closed geodesics which are more hyperbolic than the closed geodesics for the metrics of constant negative curvature with the same total area.

4(B). It is interesting to explore the relationship between the hyperbolic closed geodesics described in corollary 4.3 and the shortest geodesics in free homotopy classes from theorem 3.1. The connections do not seem to be simple. For example, I do not know whether there are always infinitely many hyperbolic closed geodesics which are the shortest in their respective free homotopy classes; even the existence of one such geodesic is not clear.

On the other hand, one can count those hyperbolic closed geodesics which are the shortest among the hyperbolic closed geodesics in respective elements of Π. In other words, let $P_\sigma^{s,h}(T)$ be the number of $\Gamma \in \Pi$ such that there is a hyperbolic closed geodesic in Γ of length $\leq T$; let, moreover,

$$P_\sigma^{s,h} = \underline{\lim}\,(\log P_\sigma^{s,h}(T))/T.$$

The following conjecture seems reasonable.

4.4. CONJECTURE. *Under the assumptions of* 4.3

$$P_\sigma^{s,h} \geq K(E, v_\sigma)$$

and this inequality is strict for every Riemannian metric σ of non-constant curvature.

4(C). In the rest of this section we will prove a refinement of corollary 4.3 which implies a weaker version of conjecture 4.4. We assume that σ is a Riemannian metric on a compact surface with negative Euler characteristic.

Let for $t > 0$

$S_{t,\sigma} = \{v \in S^\sigma M$: for every lift $\tilde{v}$ of v the geodesic $\{\pi_{\tilde{\sigma}} \phi_s^{\tilde{\sigma}} \tilde{v}\}_{s=-t}^{t}$ is one of the shortest curves connecting its ends$\}$, and

$$S_\sigma = \{v \in S^\sigma M: \text{for every lift } \tilde{v} \text{ of } v \text{ the whole geodesic } \{\pi_{\tilde\sigma}\phi_t^{\tilde\sigma}\tilde{v}\}_{t=-\infty}^{t} \text{ is the shortest curve between any two of its points}\}. \tag{4.2}$$

Obviously, if $t_1 > t_2$ then $S_{t_1,\sigma} \subset S_{t_2,\sigma}$ and

$$S_\sigma = \bigcap_{t>0} S_{t,\sigma}.$$

Every $S_{t,\sigma}$ is a closed subset of $S^\sigma M$; it is non-empty because it contains all closed geodesics of length $\geq 2t$ which are shortest in their free homotopy classes. Thus, by the compactness of M, the set S_σ is also non-empty; it is obviously closed and invariant with respect to the geodesic flow. Moreover, its dynamical structure is rich enough as follows from the next statement which is a refinement of 1.5.

4.5. Proposition. $h(\phi^\sigma|_{S_\sigma}) \geq P_\sigma^s$.

Proof. Let for a closed ϕ^σ-invariant set $A \subset S^\sigma M$, $S_\sigma(T, \delta, A)$ be the maximal number of (T, δ)-separated orbits of ϕ^σ belonging to the set A and $\hat{N}_\sigma(T, \delta, A)$ be the minimal number of sets of diameter $\leq 2\delta$ in the metric D_σ^T which cover A.

Clearly

$$\hat{N}_\sigma(T, \delta, A) \geq S_\sigma(T, 2\delta, A) \tag{4.3}$$

because every element of the covering can contain at most one point from a (T, δ)-separated set.

On the other hand

$$\hat{N}_\sigma(T_1 + T_2, \delta, A) \leq \hat{N}(T_1, \delta, A) \cdot \hat{N}(T_2, \delta, A) \tag{4.4}$$

because for any coverings $\mathcal{A}_1$ and $\mathcal{A}_2$ corresponding to the $D_\sigma^{T_1}$ and $D_\sigma^{T_2}$ one can produce a covering by sets $c \cap \phi_{-T_1}^\sigma d$, $c \in \mathcal{A}_1$, $d \in \mathcal{A}_2$, of diameter $\leq 2\delta$ with respect to $D_\sigma^{T_1+T_2}$ whose cardinality does not exceed $\operatorname{Card} \mathcal{A}_1 \cdot \operatorname{Card} \mathcal{A}_2$. It follows from (4.4) that

$$\lim_{T\to\infty} (\log \hat{N}_\sigma(T, \delta, A))/T$$

exists and is equal to

$$\inf_T (\log \hat{N}_\sigma(T, \delta, A))/T.$$

Moreover, $h(\phi^\sigma|_A) = \lim_{\delta\to 0} \lim_{T\to\infty} (\log \hat{N}(T, \delta, A))/T$.

Let

$$S'_{t,\sigma} = \bigcap_{s\in\mathbf{R}} \phi_s^\sigma S_{t,\sigma}.$$

By definition, $S'_{t,\sigma}$ is a closed invariant set of the geodesic flow; it still contains all the closed geodesics of length $\geq 2\delta$ which are the shortest in their respective free homotopy classes. Since non-homotopical closed geodesics are always separated by a constant independent of σ, say $\delta_\sigma > 0$ we obtain from (4.3) that for every t

$$\hat{N}_\sigma\left(T, \frac{\delta_\sigma}{2}, S'_{t,\sigma}\right) \geq P_\sigma^s(T). \tag{4.5}$$

Let us now fix $T>0$ and $\delta>0$. Since

$$S_\sigma=\bigcap_t S'_{t,\sigma}$$

one can find $t>0$ such that the set $S'_{t,\sigma}$ is contained in $\delta/2$-neighbourhood of S_σ in the metric D^T_σ. That means, in particular, that every covering of S_σ by the sets of diameter $\le\varepsilon$ in this metric produces a covering of $S'_{t,\sigma}$ by the same number of sets of diameter $\le\varepsilon+\delta$ (by taking δ-neighbourhoods of the elements of the covering). Thus by (4.5)

$$\hat N_\sigma\left(T,\frac{\delta_\sigma}{2}-\delta,S_\sigma\right)\ge P^s_\sigma(T)$$

so that

$$h(\phi^\sigma|_{S_\sigma})\ge\varliminf_{T\to\infty}\frac{1}{T}\left(\log\hat N_\sigma\left(T,\frac{\delta_\sigma}{4},S_\sigma\right)\right)\ge P^s_\sigma. \qquad \square$$

Our last estimate involves the counting of those hyperbolic closed geodesics which are lifted to almost shortest geodesics on the universal covering. More precisely, let for $\varepsilon>0$

$$P^\varepsilon_{\sigma,X}(T)=\text{Card}\,\{\gamma\in\mathcal{P}_{\sigma,X}(T)\text{: for every } v\in\gamma,\text{ every lift }\tilde v\text{ of }v,\ t,\ s\text{: }0\le t,\ s,\ \le l_\sigma(\gamma)$$
$$d_{\tilde\sigma}(\tilde\pi\phi^{\tilde\sigma}_t\tilde v,\ \tilde\pi\phi^{\tilde\sigma}_s\tilde v)>|t-s|+\varepsilon\}, \tag{4.6}$$

(cf. with (4.1)) and as usual

$$P^\varepsilon_{\sigma,X}=\varliminf_{T\to\infty}(\log P^\varepsilon_{\sigma,X}(T))/T.$$

4.6. THEOREM. *Under the assumptions of* 4.3 *for every* $\varepsilon>0$

$$P^\varepsilon_{\sigma,K(E,v_\sigma)}\ge K(E,v_\sigma)$$

and this inequality is strict for every metric of non-constant curvature.

The proof makes use of theorem 3.2, proposition 4.5 and theorem 4.1 with the following addition which is also a continuous time version of some information obtained in [**13**, cf. main lemma and the description of the set $L_n(\varepsilon,l)$ on p. 172].

4.7. *The statement of theorem* 4.1 *remains true if we count only those closed orbits which lie in* ε*-neighbourhood in the metric* d^f_T (*cf.* (1.3)) *of orbits belonging to the support of the measure* μ.

Proof of 4.6. We can assume that σ is not a metric of constant negative curvature because otherwise $S_\sigma=S^\sigma M$, and $P^\varepsilon_{\sigma,K(E,v_\sigma)}=P_{\sigma,K(E,v_\sigma)}=K(E,v_\sigma)$.

Let us fix $\varepsilon>0$. By proposition 4.5 and the variational principle for topological entropy [**7**] we can find a ϕ^σ-invariant ergodic measure supported on the set S_σ such that

$$h_\mu(\phi^\sigma)>h(\phi^\sigma|_{S_\sigma})=\varepsilon$$

so that if ε is sufficiently small then by theorem 3.2

$$h_\mu(\phi^\sigma)>K(E,v_\sigma).$$

By the above entropy estimate [**25**] we have for the positive Lyapunov exponent χ_2^μ of ϕ^σ with respect to μ:

$$\chi_1^\mu \geq h_\mu(\phi^\sigma) > K(E, v_\sigma).$$

Now we can apply theorem 4.1 and 4.7. If ε is small enough then for every constructed closed geodesic γ, every lift $\tilde\gamma$ of γ and $\tilde v \in \tilde\gamma$ we can find by 4.7 a tangent vector $v' \in S_\sigma$ and a lift $\tilde v'$ of v' such that for $0 \leq t \leq T$

$$D_{\tilde\sigma}(\phi_t^{\tilde\sigma}\tilde v, \phi_t^{\tilde\sigma}\tilde v') < \varepsilon/2 \tag{4.7}$$

(recall that $T \geq l_\sigma(\gamma)$).

Thus for t, s, $0 < t$, $s \leq l_\sigma(\gamma)$

$$d_{\tilde\sigma}(\tilde\pi\phi_t^{\tilde\sigma}\tilde v, \tilde\pi\phi_s^{\tilde\sigma}\tilde v) \geq d_{\tilde\sigma}(\tilde\pi\phi_t^{\tilde\sigma}\tilde v', \tilde\pi\phi_s^{\tilde\sigma}\tilde v') - d_{\tilde\sigma}(\tilde\pi\phi_t^{\tilde\sigma}\tilde v, \tilde\pi\phi_t^{\tilde\sigma}\tilde v') - d_{\tilde\sigma}(\tilde\pi\phi_s^{\tilde\sigma}\tilde v, \tilde\pi\phi_s^{\tilde\sigma}\tilde v'). \tag{4.8}$$

Since $v' \in S_\sigma$ (cf. (4.2)) the first term in the right-hand part of (4.8) is equal to the length of the geodesic segment $\{\tilde\pi\phi_t^{\tilde\sigma}\tilde v'\}_{\tau=s}^t$ i.e., $|t-s|$. Two other terms are estimated from above by ε as follows from (4.7). This means that γ satisfies (4.6). □

This work was partially supported by NSF Grant MCS79-038.

REFERENCES

[1] D. V. Anosov. Geodesic flows on closed Riemannian manifolds of negative curvature. *Proc. Steklov Inst. Math.* **90** (1967).

[2] M. Berger. Lectures on geodesics in Riemannian Geometry, Tata Institute, Bombay, 1965.

[3] M. Berger. Some relations between volume, injectivity radius and convexity radius in Riemannian manifolds. *Differential Geometry and Relativity.* Reidel: Dordrecht–Boston, 1976, 33–42.

[4] G. D. Birkhoff. Dynamical systems. *A.M.S. Colloquium Publications* 9. A.M.S.: New York, 1927.

[5] R. Bishop & R. Crittenden. *Geometry of Manifolds.* Academic Press: New York, 1964.

[6] R. Bowen. Periodic orbits for hyperbolic flows. *Amer. J. Math.* **94** (1972), 1–30.

[7] M. Denker, C. Grillenberger & K. Sigmund. Ergodic theory on compact spaces. *Lecture Notes in Math.* No 527, Springer: Berlin, 1976.

[8] E. I. Dinaburg. On the relations among various entropy characteristics of dynamical systems. *Math. USSR, Izv.* **5** (1971), 337–378.

[9] D. Gromoll, W. Klingenberg & W. Meyer. *Riemannische Geometrie in Großen.* Springer: Berlin, 1968.

[10] G. A. Hedlund. The dynamics of geodesic flows. *Bull. Amer. Math. Soc.* **45** (1939), 241–260.

[11] S. Helgason. *Differential Geometry, Lie Groups and Symmetric Spaces.* Academic Press: New York, 1978.

[12] J.A. Jenkins. *Univalent Functions and Conformal Mappings.* Springer: Berlin, 1958.

[13] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publ. Math. IHES* **51** (1980), 137–173.

[14] A. Katok. Counting closed geodesics on surfaces, Mathematische Arbeitstagung, 1980, Universität Bonn.

[15] A. Katok. Closed geodesics and ergodic theory. *London Math. Soc. Symp. on Ergodic Theory, Durham* 1980. Abstracts, University of Warwick, 1980.

[16] A. Katok. Lyapunov exponents, entropy, hyperbolic sets and ε-orbits. (In preparation.)

[17] A. Manning. Topological entropy for geodesic flows. *Ann. Math.* **110** (1979), 567–573.

[18] A. Manning. Curvature bound for the entropy of the geodesic flow on a surface. (Preprint, 1980.)

[19] G. A. Margulis. Applications of ergodic theory to the investigation of manifolds of negative curvature. *Funct. Anal. Appl.* **3** (1969), 335–336. (Translated from Russian.)

[20] G. A. Margulis. On some problems in the theory of *U*-systems *Dissertation*, Moscow State University, 1970. (In Russian.)

[21] M. Morse. The calculus of variations in the large. *A.M.S. Colloquium Publications* 18 A.M.S.: New York, 1936.
[22] M. Morse. A fundamental class of geodesics in any closed surface of genus greater than one. *Trans. A.M.S.* **26** (1924), 25–61.
[23] Ja. B. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russ. Math. Surveys* **32** (1977), 4, 55–114.
[24] H. Poincaré. Sur les lignes geodesigue des surfaces convexes. *Trans. Amer. Math. Soc.* **6** (1905), 237–274.
[25] D. Ruelle. An inequality for the entropy of differentiable map. *Bol. Soc. Bras. Mat.* **9** (1978), 83–87.
[26] P. Sarnak. Entropy estimates for geodesic flows. *Ergod. Th. & Dynam. Sys.* **2** (1982).
[27] M. Schiffer & D. C. Spencer. *Functionals on Finite Riemannian Surfaces.* Princeton Univ. Press: Princeton, 1954.
[28] A. Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Ind. Math. Soc.* **20** (1956), 47–87.
[29] Ja. G. Sinai. The asymptotic behaviour of the number of closed geodesics on a compact manifold of negative curvature. *Izv. Akad. Nauk SSSR, Ser. Math.* **30** (1966), 1275–1295. English translation, *A.M.S. Trans.* **73** 2 (1968), 229–250.

https://doi.org/10.1017/S0143385700001656 Published online by Cambridge University Press