TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 282, Number 2, April 1984

ENTROPY VIA RANDOM PERTURBATIONS

BY

YURI KIFER[1]

ABSTRACT. The entropy of a dynamical system S^t on a hyperbolic attractor with respect to the Bowen-Ruelle-Sinaĭ measure is obtained as a limit of entropy characteristics of small random perturbations x_t^ε of S^t. Both the case of perturbations only in some neighborhood of an attractor and global perturbations of a flow with hyperbolic attracting sets are considered.

1. Introduction. Let M be a locally compact, separable, connected d-dimensional C^3-class Riemannian manifold, and let $B, X_1, X_2, \ldots, X_d$ be C^2-vector fields on M. Consider the flow S^t: $M \to M$, $-\infty < t < \infty$, defined by

$$d(S^t x)/dt = B(S^t x), \tag{1.1}$$

and the family of diffusion Markov processes x_t^ε given by the stochastic differential equation

$$dx_t^\varepsilon = B(x_t^\varepsilon)\,dt + \varepsilon \sum_{1 \le i \le d} X_i(x_t^\varepsilon)\,dw_t^i, \tag{1.2}$$

where $w_t = (w_t^1, \ldots, w_t^d)$ is a standard d-dimensional Brownian motion, $\varepsilon > 0$ is a parameter, and the stochastic differential in (1.2) is taken in the sense of Stratonovitch (see [**6**]) so that its infinitesimal generator L^ε has the form

$$L^\varepsilon = B + \frac{\varepsilon^2}{2} \sum_{1 \le i \le d} X_i X_i. \tag{1.3}$$

One can understand (1.2) in the sense that for any smooth function f on M,

$$f(x_t^\varepsilon) = f(x_0^\varepsilon) + \int_0^t Bf(x_s^\varepsilon)\,ds + \varepsilon \sum_{1 \le i \le d} \int_0^t X_i f(x_t^\varepsilon)\,dw_t^i. \tag{1.4}$$

The processes x_t^ε are called small random perturbations of the flow S^t. This definition is essentially the same as in [**8** and **9**], but the present notation is more convenient for our goals and also for clarifying this model of random perturbations. One can consider (1.2) as the equation (1.1) with the right-hand side of it perturbed by random vector fields. The representation of x_t^ε as a stochastic flow (see [**6**]) is also relevant here.

Received by the editors February 9, 1983.
1980 *Mathematics Subject Classification*. Primary 58F30, 58G32, 58F15.
Key words and phrases. Hyperbolic attractor, diffusion process, entropy.
[1]Supported by U.S.A.–Israel BSF.

Let $\Gamma = (V_1, \ldots, V_k)$ be a partition of M into closed subsets such that

$$\bigcup_{1 \leq i \leq k} V_i = M \quad \text{and} \quad V_i \cap V_j = \partial V_i \cap \partial V_j \quad \text{if } i \neq j, \tag{1.5}$$

where $\partial V_i = V_i \backslash \operatorname{int} V_i$ is the boundary of V_i. Denote

$$P^\varepsilon(t, x, Q) = P_x^\varepsilon\{x_t^\varepsilon \in Q\}, \tag{1.6}$$

for any Borel subset Q, where $P_x^\varepsilon\{\cdot\}$ is the probability of the event in brackets, provided $x_0^\varepsilon = x$. Suppose x_t^ε has an invariant probability measure μ^ε, i.e.

$$\mu^\varepsilon(Q) = \int_M \mu^\varepsilon(dx) P^\varepsilon(t, x, Q) \tag{1.7}$$

for any Borel subset Q and $t > 0$. Then one can consider

$$\sigma^\varepsilon(i_0, \ldots, i_n) = \int_{V_{i_0}} \mu^\varepsilon(dx) P_x^\varepsilon\{x_1^\varepsilon \in V_{i_1}, x_2^\varepsilon \in V_{i_2}, \ldots, x_n^\varepsilon \in V_{i_n}\}. \tag{1.8}$$

Assume

$$\mu^\varepsilon(\partial\Gamma) = 0, \tag{1.9}$$

where $\partial\Gamma = \bigcup_{i=1}^k \partial V_i$ and set

$$H_n^\varepsilon(\Gamma) = -\sum_{i_0, \ldots, i_n} \sigma^\varepsilon(i_0, \ldots, i_n) \ln \sigma^\varepsilon(i_0, \ldots, i_n), \tag{1.10}$$

where $0 \ln 0 \equiv 0$. It is known (see [**2**]) that

$$H_{N+n}^\varepsilon(\Gamma) \leq H_N^\varepsilon(\Gamma) + H_n^\varepsilon(\Gamma), \tag{1.11}$$

so by the standard subadditivity argument the limit

$$H^\varepsilon(\Gamma) = \lim_{N \to \infty} \frac{1}{N+1} H_N^\varepsilon(\Gamma) = \inf_{N \geq 0} \frac{1}{N+1} H_N^\varepsilon(\Gamma) \tag{1.12}$$

exists. The quantity $H^\varepsilon(\Gamma)$ is the entropy (see [**2**]) of the stationary Markov chain x_n^ε, $n = 0, 1, \ldots,$ with the invariant measure μ^ε with respect to the partition of the sample space Ω into the sets

$$\Omega_i = \{\omega : x_0^\varepsilon(\omega) \in V_i\}, \qquad i = 1, \ldots, k.$$

Assume

$$\mu^{\varepsilon_j} \to \mu \quad \text{as } \varepsilon_j \to 0 \text{ in the weak sense} \qquad \left(\mu^{\varepsilon_j} \xrightarrow{\text{w}} \mu\right) \tag{1.13}$$

and

$$\lim_{\varepsilon \to 0} \sup_x \left| \int_M P^\varepsilon(t, x, dy) f(y) - f(S^t x) \right| = 0 \tag{1.14}$$

for any $t > 0$ and a continuous function f on M. Then (see [**5**]) the measure μ will be invariant for the flow S^t. If

$$\mu(\partial\Gamma) = 0 \tag{1.15}$$

then one can also consider the entropy $H(\Gamma)$ of the transformation S^1 with respect to the partition Γ (see [**2**]),

$$H(\Gamma) = \lim_{N\to\infty} \frac{1}{N+1} H_N(\Gamma) = \inf_{N\geqslant 0} \frac{1}{N+1} H_N(\Gamma), \tag{1.16}$$

where

$$H_n(\Gamma) = -\sum_{i_0,\dots,i_N} \sigma(i_0,\dots,i_N) \ln \sigma(i_0,\dots,i_N)$$

(again $0 \ln 0 \equiv 0$) and

$$\sigma(i_0,\dots,i_N) = \mu\Big(\bigcap_{0\leqslant j\leqslant N} S^{-j} V_{i_j}\Big). \tag{1.17}$$

From (1.9) and (1.13)–(1.17) it follows that

$$\sigma^{\varepsilon_j}(i_0,\dots,i_N) \to \sigma(i_0,\dots,i_N) \quad \text{as } \varepsilon_j \to 0, \tag{1.18}$$

so by (1.12), (1.15) and (1.16),

$$\limsup_{\varepsilon_j\to 0} H^{\varepsilon_j}(\Gamma) \leqslant \inf_N \frac{1}{N+1} \limsup_{\varepsilon_j\to 0} H_N^{\varepsilon_j}(\Gamma) = H(\Gamma). \tag{1.19}$$

We shall prove in this paper that if S^t is a hyperbolic flow, Γ is a partition into elements of small size satisfying (1.9) and (1.14), and μ is the Sinaĭ-Bowen-Ruelle measure on an attractor, then

$$\lim_{\varepsilon\to 0} H^{\varepsilon}(\Gamma) = h_\mu(S^1), \tag{1.20}$$

where $h_\mu(S^1) = \sup_\Gamma H(\Gamma)$ is the entropy of S^1 with respect to the measure μ (see [**2**]).

Notice that in the case $h_\mu(S^1) = 0$, the inequality (1.19) automatically becomes an equality, so (1.20) is also true. Relation (1.20) is interesting also in the framework of the general question: What parameters of dynamical systems can be approximated by parameters of their random perturbations?

We shall consider two cases. The first is the case of perturbations considered only in some neighborhood of a hyperbolic attractor Λ, i.e. when the operator L^ε coincides with B outside of some neighborhood of Λ and L^ε is a nondegenerate elliptic operator in some smaller neighborhood of Λ. This case is important since a hyperbolic attractor can be constructed in a noncompact space, such as R^d, and it is not natural to immerse it first into a compact manifold and only then to consider random perturbations on this manifold.

We also prove (1.20) in the second case when M is a compact manifold, x_t^ε is a nondegenerate diffusion on M, and the ω-limit set of S^t is the disjoint union of a finite number of hyperbolic attractors and a finite number of closed unstable sets. The measure μ in this case is a combination of Sinaĭ-Bowen-Ruelle measures on attractors. As usual, by simplifying our arguments one can reproduce all results for random perturbations (of the type considered in [**8**]) of hyperbolic diffeomorphisms.

Notice that under natural regularity conditions on the transition probability $P^\varepsilon(t, x, Q)$, such as Döblin's condition (see [**4**]), which are satisfied provided the

operator L^ε is nondegenerate and M is compact, the entropy of the Markov chain x_n^ε, i.e. $\sup_\Gamma H^\varepsilon(\Gamma)$, is infinite. Therefore this notion is not interesting here. Nevertheless, the entropy with respect to partitions and the corresponding asymptotics turn out to be very useful for small random perturbations, since we can first let $\varepsilon \to 0$ and then, if necessary, take the supremum over partitions.

The structure of this paper is the following. In §2 we give the definitions and formulate the main results. §§3 and 4 contain the proofs. In the conclusion (§5) we consider another approach and discuss the case of a nonuniform hyperbolic flow (see [**11**]).

This paper was written during my visit to the University of Maryland. The author gratefully acknowledges the University of Maryland for the invitation to spend my sabbatical there and A. Katok for a number of very useful discussions.

2. Definitions and main results. A compact S^t-invariant $(-\infty < t < \infty)$ set $\Lambda \subset M$ is called hyperbolic if the tangent bundle TM restricted to Λ can be written as the Whitney sum of continuous subbundles

$$T_\Lambda M = F^s \oplus F^0 \oplus F^u, \tag{2.1}$$

where $F^0 = B$, this decomposition is invariant with respect to the differential DS^t of S^t, and there exist $c, \nu > 0$ such that

$$\|DS^t\xi\| \leqslant ce^{-\nu t}\|\xi\| \quad \text{for } \xi \in F^s,\ t \geqslant 0, \tag{2.2}$$

and

$$\|DS^{-t}\eta\| \leqslant ce^{-\nu t}\|\eta\| \quad \text{for } \eta \in F^u,\ t \geqslant 0. \tag{2.3}$$

A set Λ is called a hyperbolic (basic) attractor if:

(a) Λ is hyperbolic and either Λ is a point or it contains no fixed points;

(b) the periodic orbits of $S^t|_\Lambda$ are dense in Λ;

(c) $S^t|_\Lambda$ is a topologically transitive flow;

(d) there is an open set $U_\Lambda \supset \Lambda$ such that

$$S^tU_\Lambda \subset U_\Lambda \quad \text{for all } t > 0 \quad \text{and} \quad \Lambda = \bigcap_{t \geqslant 0} S^tU_\Lambda. \tag{2.4}$$

For $x \in \Lambda$ let $J_t(x)$ be the Jacobian of the linear map $DS^t\colon F_x^u \to F_{S^tx}^u$. Define

$$\varphi^{(u)}(x) = -\frac{dJ_t(x)}{dt}\bigg|_{t=0}. \tag{2.5}$$

If Λ is a hyperbolic attractor then there exists a unique S^t-invariant probability measure μ_Λ on Λ such that the entropy $h_{\mu_\Lambda}(S^1)$ of S^1 with respect to μ_Λ satisfies the formula (see [**3**])

$$h_{\mu_\Lambda}(S^1) = -\int_\Lambda \varphi^{(u)}(x)\mu_\Lambda(dx). \tag{2.6}$$

ASSUMPTION A. *Let Λ be a hyperbolic attractor, an open set U_Λ has a compact closure and satisfies* (2.4), *the operator L^ε defined in* (1.3) *is a nondegenerate elliptic operator in some open set $U_L \subset U_\Lambda$ and $L^\varepsilon \equiv B|_{M \setminus U_\Lambda}$.*

If Assumption A is satisfied then it follows that the process x_t^ε never leaves $\overline{U}_\Lambda \equiv U_\Lambda \cup \partial U_\Lambda$ provided $x_0^\varepsilon \in \overline{U}_\Lambda$. Using [**8**] we shall prove the following result in §3.

PROPOSITION 1. *Suppose Assumption* A *is valid. Then the process* x_t^ε *has the unique invariant measure* μ_A^ε *having support in* $\overline{U}_\Lambda$ *and*

$$\mu_A^\varepsilon \xrightarrow{w} \mu_\Lambda \quad \text{as } \varepsilon \to 0. \tag{2.7}$$

The main result of this paper is the following

THEOREM 1. *Let Assumption* A *be satisfied and* $\Gamma = (V_1, \dots, V_k)$ *be a partition of* M *such that*

$$\partial\Gamma \subset U_L, \qquad m(\partial\Gamma) = \mu_\Lambda(\partial\Gamma) = 0 \tag{2.8}$$

and $\max_{V_i \in \Gamma_\Lambda(\rho)} \operatorname{diam} V_i \leqslant \rho$, *where* $\Gamma_\Lambda(\rho) \equiv \{V_i\colon V_i \cap U_\rho(\Lambda) \neq \varnothing\}$, $U_\rho(\Lambda) \equiv \{y\colon \operatorname{dist}(y, \Lambda) \leqslant \rho\}$, $\rho > 0$ *is small enough and* m *is the Riemannian volume.*

Then

$$\lim_{\varepsilon \to 0} H^\varepsilon(\Gamma) = h_{\mu_\Lambda}(S^1), \tag{2.9}$$

where $H^\varepsilon(\Gamma)$ *is defined by* (1.8), (1.10) *and* (1.12) *with* $\mu^\varepsilon = \mu_A^\varepsilon$ *given by Proposition* 1.

REMARK 1. In (2.9) one can also take a partition Γ with piecewise smooth boundary $\partial\Gamma \subset U_L$. In this case (2.8) will be satisfied and the left-hand side of (2.9) will not depend on the dynamical system S^t but only on the process x_t^ε.

For any $x, y \in M$ define

$$d(x, y) = \inf_{0 \leqslant \tau \leqslant 1} \inf_{x_0 = S^\tau x,\, x_n = y,\, n \geqslant 0} \sum_{0 \leqslant i \leqslant n-1} \operatorname{dist}\left(S^1 x_i, x_{i+1}\right)^2, \tag{2.10}$$

where inf is taken over all finite sequences $\{x_i \in M, 1 \geqslant \tau \geqslant 0, i = 0, \dots, n;\ x_0 = S^\tau x$ and $x_n = y\}$. Following [**12**] we shall say that x is equivalent to y ($x \sim y$) if $d(x, y) = 0$ and $d(y, x) = 0$. A set $K \subset M$ is called stable if for any $x \in K$ and $y \notin K$ one has $d(x, y) > 0$.

ASSUMPTION B. (a) *M is a compact manifold and the operator L^ε defined by* (1.3) *is a nondegenerate elliptic operator on M*; (b) *there exist a finite number of equivalence classes $K_1, \dots, K_l$ (with respect to the relation $\sim$ above) which are compact sets and contain any ω-limit set of the equation* (1.1); (c) *the stable sets among $K_1, \dots, K_l$ are $K_{i_1} = \Lambda_1, \dots, K_{i_\nu} = \Lambda_\nu$ and they are hyperbolic attractors.*

REMARK 2. Axiom A flows (see [**3**]) satisfy Assumption B.

From (a) of Assumption B it follows that the transition probability $P^\varepsilon(t, x, Q)$ of the process x^ε has a positive density with respect to the Riemannian volume m and so satisfies the Döblin condition (see [**4**]). Therefore it has a unique invariant measure μ_B^ε. It follows from [**8** and **12**] under conditions of Assumption B that if

$$\mu_B^{\varepsilon_i} \xrightarrow{w} \mu_0 \quad \text{as } \varepsilon_i \to 0, \tag{2.11}$$

then

$$\mu_0 = \sum_{1 \leqslant j \leqslant \nu} q_j \mu_{\Lambda_j}, \tag{2.12}$$

where $q_j \geqslant 0$, $\sum_{1 \leqslant j \leqslant \nu} q_j = 1$ and $\{\mu_{\Lambda_j}\}$ are defined by (2.6) on the hyperbolic attractors $\{\Lambda_j\}$.

THEOREM 2. *Suppose Assumption* B *and* (2.11) *are satisfied. Let* $\Gamma = (V_1, \dots, V_k)$ *be a partition of* M *such that*

$$m(\partial\Gamma) = \mu_{\Lambda_i}(\partial\Gamma) = 0 \quad \textit{for all } i = 1, \dots, \nu \tag{2.13}$$

and

$$\max_i \max_{V_j \in \Gamma_{\Lambda_i}} \operatorname{diam} V_j \leqslant \rho, \tag{2.14}$$

where $\rho > 0$ *is small enough and* Γ_{Λ_i} *is defined in the same way as in Theorem* 1. *Then*

$$\lim_{\varepsilon_i \to 0} H^{\varepsilon_i}(\Gamma) = \sum_{1 \leqslant j \leqslant \nu} h_{\mu_{\Lambda_j}}(S^1), \tag{2.15}$$

where $\{\varepsilon_i\}$ *is the same sequence as in* (2.11) *and* $H^\varepsilon(\Gamma)$ *is defined by* (1.8), (1.10) *and* (1.12) *with* $\mu^\varepsilon = \mu^\varepsilon_B$.

Remark 1 is relevant to this case, as well.

REMARK 3. It is easy to see that (1.14) is satisfied under either of Assumptions A or B.

3. Proof of Proposition 1. Let U_0, U_1 and U_2 be open sets with smooth boundaries U_0, U_1, and ∂U_2 such that

$$\Lambda \subset U_0 \subset U_0 \cup \partial U_0 \subset U_1 \subset U_1 \cup \partial U_1 \subset U_2 \subset U_2 \cup \partial U_2 \subset U_L. \tag{3.1}$$

Suppose that Assumption A is satisfied. Then $x^\varepsilon_t \in U_\Lambda \cup \partial U_\Lambda$ for all $t \geqslant 0$ provided $x^\varepsilon_0 \in U_\Lambda \cup \partial U_\Lambda$.

It is clear that there is $T > 0$ such that for any $x \in U_\Lambda \cup \partial U_\Lambda$ it follows that $S^T x \in U_0$. Then one can see that

$$q^\varepsilon = \inf_{x \in U_\Lambda} P^\varepsilon(T, x, U_1) \to 1 \quad \text{as } \varepsilon \to 0. \tag{3.2}$$

Let $\tau_0 = 0$, let τ_1 be the first time the process x^ε_t hits ∂U_1, and let $\tilde\tau_1$ be the first time after τ_1 that x^ε_t hits ∂U_2. Let $\tau_i(\tilde\tau_i)$ be the first time after $\tilde\tau_{i-1}(\tau_i)$ that x^ε_t hits ∂U_1 (∂U_2).

By the strong Markov property, for any $x \in U_1 \cup \partial U_1$ and any Borel set $V \subset U_0$ one has

$$\begin{aligned} P^\varepsilon(t, x, V) &= \tilde P^\varepsilon(t, x, V) + E^\varepsilon_x \chi_{\tau_2 < t} P^\varepsilon\big(t - \tau_2, x^\varepsilon_{\tau_2}, V\big) \\ &= \sum_{n=0}^{\infty} E^\varepsilon_x \chi_{\tau_n < t} \tilde P^\varepsilon\big(t - \tau_n, x^\varepsilon_{\tau_n}, V\big), \end{aligned} \tag{3.3}$$

where E^ε_x is the expectation of the process x^ε_t starting at x, $\chi_A = 1$ if an event A occurs and $\chi_A = 0$ otherwise, $\tilde P^\varepsilon(s, x, V)$ is the transition probability of the process x^ε_t with the absorption on ∂U_2, i.e. the solution of the parabolic equation $\partial u^\varepsilon/\partial t = L^\varepsilon u^\varepsilon$ with zero conditions on ∂U_2.

Since L^ε is uniformly nondegenerate in U_2 (while $\varepsilon > 0$ is fixed) then (see [**1**]) there exist $\tilde{\tilde q}^\varepsilon > \tilde q^\varepsilon > 0$ such that

$$\tilde{\tilde q}^\varepsilon \geqslant \tilde p^\varepsilon(1, x, y) \geqslant \tilde q^\varepsilon \quad \text{for any } x, y \in U_1 \cup \partial U_1 \tag{3.4}$$

and

$$\tilde{p}^{\epsilon}(t, x, y) \leqslant \tilde{\tilde{q}}^{\epsilon} \quad \text{for any } t \geqslant 0 \text{ if } x \in \partial U_1 \text{ and } y \in U_0, \tag{3.5}$$

where $\tilde{p}^{\epsilon}(t, x, y)$ is the transition density of the process x_t^{ϵ} with absorption on ∂U_2, i.e.

$$\tilde{P}^{\epsilon}(t, x, \tilde{V}) = \int_{\tilde{V}} \tilde{p}^{\epsilon}(t, x, y) m(dy) \tag{3.6}$$

for any Borel set V.

From the strong Markov property and (3.2) it follows that

$$\sup_{x \in U_\Lambda} P_x^{\epsilon}\{\tau_n < T\} \leqslant (1 - q^{\epsilon})^n. \tag{3.7}$$

Therefore by (3.3)–(3.7),

$$(1 + 1/q^{\epsilon})\tilde{\tilde{q}}^{\epsilon} m(V) \geqslant P^{\epsilon}(T + 1, x, V) \geqslant q^{\epsilon} \cdot \tilde{q}^{\epsilon} m(V). \tag{3.8}$$

Since U_0, U_1, and U_2 satisfying (3.1) can be arbitrary, it follows from (3.8) that $P^{\epsilon}(T + 1, x, V)$ has a positive density $p^{\epsilon}(T + 1, x, y)$ for any $x \in U_\Lambda \cup \partial U_\Lambda$ and $y \in U_L$, which is bounded away from infinity and zero provided y is in a compact subset of U_L.

Since by the Chapman-Kolmogorov formula,

$$P^{\epsilon}(t, x, V) = \int_{\overline{U}_\Lambda} P^{\epsilon}(t - T - 1, x, dy) P^{\epsilon}(T + 1, y, V), \tag{3.9}$$

then $P^{\epsilon}(t, x, V)$ has a positive density $p^{\epsilon}(t, x, y)$ for any $t \geqslant T + 1$.

Therefore the process x_t^{ϵ} satisfies the Döblin condition (see [**4**, Chapter 5]), so it has a unique invariant measure μ_A^{ϵ} with support in $\overline{U}_\Lambda$. Moreover, μ_A^{ϵ} has a positive density $g_A^{\epsilon}(x)$ in U_L with respect to the Riemannian volume m and this density is bounded away from zero and infinity in any compact subset of U_L. Hence, under the conditions of Theorem 1, when $m(\partial\Gamma) = 0$ and $\partial\Gamma \subset U_L$, one obtains

$$\mu_A^{\epsilon}(\partial\Gamma) = 0. \tag{3.10}$$

The reduction of the transition probabilities $P^{\epsilon}(t, x, V)$ in (3.3) to transition probabilities of the process x_t^{ϵ} with absorption on ∂U_2 enables us to consider x_t^{ϵ} in the domain U_2 only, where it is a nondegenerate diffusion. Then one can employ methods of [**8**] to prove (2.7) since the proof in [**8**] uses only local behaviour of x_t^{ϵ} near Λ. This remark completes the proof of Proposition 1.

4. Proofs of Theorems 1 and 2. Let

$$d_T(x, y) = \sup_{0 \leqslant t \leqslant T} \operatorname{dist}(S^t x, S^t y) \tag{4.1}$$

and

$$B_x(\delta, T) = \{y \in M : d_T(x, y) \leqslant \delta\}. \tag{4.2}$$

According to [**3**] a subset $E \subset \Lambda$ is called (δ, T)-separated if whenever $x, y \in E$, $x \neq y$, it follows that $d_T(x, y) > \delta$. It is clear that if E is a (δ, T)-separated set and $y, z \in E, y \neq z$, then $B_y(\delta/2, T)$ and $B_z(\delta/2, T)$ are disjoint, so E is a finite set and the number of points of E does not exceed some constant $K(\delta, T)$ depending only

on δ and T. Thus there exists a maximal (δ, T)-separated set $E(\delta, T)$ (not necessarily unique).

If $E(\delta, T)$ is a maximal (δ, T)-separated set of a hyperbolic attractor Λ, then (see [3])

$$\bigcup_{x \in E(\delta, T)} B_x(2\delta, T) \supset \bigcup_{x \in \Lambda} B_x(\delta, T) \supset U_{\alpha_1 \cdot \delta}(\Lambda) \supset \Lambda, \tag{4.3}$$

where $\alpha_1 > 0$ is independent of δ, T.

Since μ_Λ is ergodic (see [3]), one has from (2.5) and (2.6),

$$\lim_{T \to \infty} \frac{1}{T} \ln J_T(x) = h_{\mu_\Lambda}(S^1) \quad \text{for } \mu_\Lambda\text{-almost all } x \in \Lambda. \tag{4.4}$$

Define

$$Q_\gamma(T) = \left\{ x \in \Lambda : \left| \frac{1}{T} \ln J_T(x) - h_{\mu_\Lambda}(S^1) \right| > \gamma \right\}. \tag{4.5}$$

From [3] it follows that there exist $C_\delta > 0$ independent of T and $x \in \Lambda$ such that if δ is small enough, then

$$m(B_x(\delta, T)) \leqslant C_\delta \mu_\Lambda(B_x(\delta, T)), \tag{4.6}$$

$$C_\delta^{-1} \leqslant m(B_x(\delta, T)) J_T(x) \leqslant C_\delta, \tag{4.7}$$

and for any $y \in B_x(\delta, T) \cap \Lambda$,

$$C_\delta^{-1} \leqslant J_T(x) J_T^{-1}(y) \leqslant C_\delta. \tag{4.8}$$

If $x \in Q_{2\gamma}(T)$ and

$$T > (1/\gamma) \ln C_\delta, \tag{4.9}$$

then by (4.8),

$$B_x(\delta, T) \cap \Lambda \subset Q_\gamma(T). \tag{4.10}$$

Now let $E(\delta, T)$ be a maximal (δ, T)-separated set. Then by (4.5)–(4.10),

$$\begin{aligned} \mu_\Lambda(Q_\gamma(T)) &\geqslant \mu_\Lambda\Bigg(\bigcup_{x \in E(\delta, T) \cap Q_{2\gamma}(T)} B_x(\delta, T) \Bigg) \\ &\geqslant C_{\delta/2}^{-2} \sum_{x \in E(\delta, T) \cap Q_{2\gamma}(T)} J_T^{-1}(x). \end{aligned} \tag{4.11}$$

Now let $\Gamma = \{V_1, \ldots, V_k\}$ be a partition of M satisfying conditions of Theorem 1, and let $\Gamma_\Lambda(\rho) = \{V_i : V_i \cap U_\rho(\Lambda) \neq \varnothing\}$ such that

$$\max_{V_i \in \Gamma_\Lambda} \operatorname{diam} V_i \leqslant \rho \leqslant \frac{1}{3} \alpha_1 \delta \leqslant \frac{1}{3} \delta. \tag{4.12}$$

Define $U_\Lambda^\Gamma = \bigcup_{V_i \in \Gamma_\Lambda} V_i$. Then

$$\Lambda \subset U_\Lambda^\Gamma \subset U_{\alpha_1 \delta}(\Lambda). \tag{4.13}$$

Choose $\beta > 0$ small enough and set

$$n(\varepsilon) = \text{integral part of } \varepsilon^{-\beta}. \tag{4.14}$$

By the Markov property

$$(4.15)\quad \sigma^{\varepsilon}(i_0,\dots,i_{N\cdot n(\varepsilon)}) \leqslant \sup_z P_z^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_1},\dots,x_{n(\varepsilon)}^{\varepsilon}\in V_{i_{n(\varepsilon)}}\}$$
$$\times \sup_z P_z^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_{n(\varepsilon)+1}},\dots,x_{2n(\varepsilon)}^{\varepsilon}\in V_{2n(\varepsilon)}\}$$
$$\cdots \sup_z P_x^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_{(N-1)n(\varepsilon)+1}},\dots,x_{N\cdot n(\varepsilon)}^{\varepsilon}\in V_{i_{N\cdot n(\varepsilon)}}\}.$$

where the sup is taken over $z\in\overline{U}_\Lambda$. Therefore

(4.16)

$$H_{N\cdot n(\varepsilon)}^{\varepsilon}(\Gamma) = -\sum_{i_0,\dots,i_{N\cdot n(\varepsilon)}} \sigma^{\varepsilon}(i_0,\dots,i_{N\cdot n(\varepsilon)})\ln\sigma^{\varepsilon}(i_0,\dots,i_{N\cdot n(\varepsilon)})$$
$$\geqslant -\sum_{k=1}^{N}\Bigg(\sum_{i_{(k-1)n(\varepsilon)+1},\dots,i_{kn(\varepsilon)}} \sigma^{\varepsilon}(i_{(k-1)n(\varepsilon)+1},\dots,i_{kn(\varepsilon)})$$
$$\times\ln\sup_z P_x^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_{(k-1)n(\varepsilon)+1}},\dots,x_{n(\varepsilon)}^{\varepsilon}\in V_{i_{kn(\varepsilon)}}\}\Bigg)$$
$$= -N\sum_{i_1,\dots,i_{n(\varepsilon)}} \sigma^{\varepsilon}(i_1,\dots,i_{n(\varepsilon)})\ln\sup_z P_z^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_1},\dots,x_{n(\varepsilon)}^{\varepsilon}\in V_{i_{n(\varepsilon)}}\}$$
$$\overset{\text{def}}{\equiv} N\tilde{H}_{n(\varepsilon)}^{\varepsilon}(\Gamma).$$

Moreover,

$$(4.17)\quad \tilde{H}_{n(\varepsilon)}^{\varepsilon}(\Gamma) \geqslant \mathcal{H}^{\varepsilon}(\mathcal{F}_{n(\varepsilon)}^{\gamma})$$
$$\overset{\text{def}}{\equiv} \sum_{(i_1,\dots,i_{n(\varepsilon)})\in\mathcal{F}_{n(\varepsilon)}^{\gamma}} \sigma^{\varepsilon}(i_1,\dots,i_{n(\varepsilon)})\ln\sup_z P_z^{\varepsilon}\{x_1^{\varepsilon}\in V_{i_1},\dots,x_{n(\varepsilon)}^{\varepsilon}\in V_{i_{n(\varepsilon)}}\},$$

where $\mathcal{F}_{n(\varepsilon)}^{\gamma}$ is the collection of $n(\varepsilon)$-sequences $i_1,\dots,i_{n(\varepsilon)}$ such that if $(i_1,\dots,i_{n(\varepsilon)})\in\mathcal{F}_{n(\varepsilon)}^{\gamma}$, then

(i) $V_{i_j}\cap U_\rho(\Lambda)\neq\varnothing$, i.e. $V_{i_j}\in\Gamma_\Lambda(\rho)$ for all $j=1,\dots,n(\varepsilon)$;

(ii) there exists $x\in E(\delta, n,(\varepsilon))$ such that $V_{i_j}\subset U_{2\delta}(S^jx)$, $j=1,\dots,n(\varepsilon)$, where $E(\delta, n(\varepsilon))$ is a maximal $(\delta, n(\varepsilon))$-separated set:

(iii) the point x from (ii) does not belong to $Q_{2\gamma}(n(\varepsilon))$.

It follows from [**8**, **9** and **12**] that

$$(4.18)\qquad \mu_A^{\varepsilon}(\overline{U}_\Lambda\backslash U_\Lambda^{\Gamma}) \leqslant C^{(1)}\exp(-\alpha_2/\varepsilon^2),$$

and for any $z\in\overline{U}_\Lambda$,

$$(4.19)\qquad P_z^{\varepsilon}\Big\{\inf_{y\in U_\rho(z)}\sup_{0\leqslant t\leqslant n(\varepsilon)}\operatorname{dist}(x_t^{\varepsilon}, S^ty)\geqslant\rho\Big\} \leqslant C^{(1)}\exp\Big(\frac{-\alpha_2}{\varepsilon}\Big)$$

for some $C^{(1)}$, $\alpha_2>0$ independent of ε, provided β in (4.14) is small enough.

On the other hand, if δ is small enough, then by the arguments of [**8**] and §5 of [**9**] for any $x\in\Lambda$ and $y\in U_\Lambda$ one can see that

$$(4.20)\qquad P_y^{\varepsilon}\Big\{\max_{1\leqslant j\leqslant n(\varepsilon)}\operatorname{dist}(x_j^{\varepsilon}, S^jx)\leqslant 2\delta\Big\} \leqslant C^{(2)}\varepsilon^{-d_u}J_{n(\varepsilon)}^{-1}(x)$$

and

$$\int_{U_\Lambda} \mu^\varepsilon(dy) P_y^\varepsilon\Big\{ \max_{1\le j\le n(\varepsilon)} \operatorname{dist}\big(x_j^\varepsilon, S^j x\big) \le 2\delta \Big\} \le C^{(2)} J_{n(\varepsilon)}^{-1}(x) \tag{4.21}$$

for some $C^{(2)} > 0$ independent of ε, x and y, where d_u is the dimension of the unstable subbundle F^u.

From (4.20), together with (i)–(iii), it follows that

$$\begin{aligned} \frac{1}{n(\varepsilon)} \mathcal{H}^\varepsilon\big(\mathcal{F}_{n(\varepsilon)}^\gamma\big) \ge \Big(h_{\mu_\Lambda}(S^1) + \frac{d_u}{n(\varepsilon)} \ln \varepsilon - 2\gamma - \frac{1}{n(\varepsilon)} \ln C^{(2)} \Big) \\ \times \sum_{(i_1,\dots,i_{n(\varepsilon)}) \in \mathcal{F}_{n(\varepsilon)}^\gamma} \sigma^\varepsilon(i_1,\dots,i_{n(\varepsilon)}). \end{aligned} \tag{4.22}$$

Next from (4.3), (4.11), (4.12), (4.18), (4.19) and (4.21), one obtains

$$\begin{aligned} &\sum_{(i_1,\dots,i_{n(\varepsilon)}) \notin \mathcal{F}_{n(\varepsilon)}^\gamma} \sigma^\varepsilon(i_1,\dots,i_{n(\varepsilon)}) \\ &\quad \le n(\varepsilon) \mu_A^\varepsilon\big(\overline{U}_\Lambda \backslash U_\Lambda^\Gamma\big) + \sup_z P_z^\varepsilon\Big\{ \inf_{y\in U_\rho(z)} \sup_{0\le t\le n(\varepsilon)} \operatorname{dist}\big(x_t, S^t y\big) \ge \rho \Big\} \\ &\qquad + \sum_{x\in E(\delta,T)\cap Q_{2\gamma}(n(\varepsilon))} \int_{U_\Lambda} \mu^\varepsilon(dy) P_y^\varepsilon\Big\{ \max_{1\le j\le n(\varepsilon)} \operatorname{dist}\big(x_j^\varepsilon, S^j x\big) \le 2\delta \Big\} \\ &\quad \le (n(\varepsilon)+1)\exp(-\alpha_2/\varepsilon) + C^{(2)} C_{\delta/2} \mu_\Lambda\big(Q_\gamma(n(\varepsilon))\big). \end{aligned} \tag{4.23}$$

From (4.4) it follows that

$$\lim_{\varepsilon\to 0} \mu_\Lambda\big(Q_\gamma(n(\varepsilon))\big) = 0. \tag{4.24}$$

By (3.10),

$$\sum_{i_1,\dots,i_{n(\varepsilon)}} \sigma^\varepsilon(i_1,\dots,i_{n(\varepsilon)}) = 1,$$

so (4.22)–(4.24) yield

$$\liminf_{\varepsilon\to 0} \frac{1}{n(\varepsilon)} \mathcal{H}^\varepsilon\big(\mathcal{F}_{n(\varepsilon)}^\gamma\big) \ge h_{\mu_\Lambda}(S^1) - 2\gamma. \tag{4.25}$$

Since γ can be arbitrarily small, one obtains from here and (4.16)–(4.17) that

$$\liminf_{\varepsilon\to 0} \frac{1}{n(\varepsilon)} \tilde{H}_{n(\varepsilon)}^\varepsilon(\Gamma) \ge h_{\mu_\Lambda}(S^1). \tag{4.26}$$

On the other hand, the limit (1.12) exists, so by (4.16),

$$H^\varepsilon(\Gamma) \ge (1/n(\varepsilon)) \tilde{H}_{n(\varepsilon)}^\varepsilon(\Gamma), \tag{4.27}$$

and by (4.26),

$$\liminf_{\varepsilon\to 0} H^\varepsilon(\Gamma) \ge h_{\mu_\Lambda}(S^1). \tag{4.28}$$

This together with (1.19) gives (2.9), since in our circumstances (2.7) is valid, so in (1.13) and (1.19) one can take any subsequence. Theorem 1 is proved.

Now suppose the conditions of Theorem 2 are satisfied and $\Gamma = \{V_1, \dots, V_k\}$ is a corresponding partition of M. Then by (1.19), (2.11) and (2.12),

$$\limsup_{\varepsilon_i \to 0} H^{\varepsilon_i}(\Gamma) \leqslant H(\Gamma) \leqslant h_{\mu_0}(S^1) = \sum_{1 \leqslant j \leqslant \nu} q_j h_{\mu_{\Lambda_j}}(S^1). \tag{4.29}$$

Inequality (4.16) does not use any assumption of Theorem 1, so this inequality remains valid in the circumstances of Theorem 2. We shall use the same notation as in (4.16).

$$\begin{aligned} \tilde{H}^{\varepsilon}_{n(\varepsilon)}(\Gamma) \equiv & - \sum_{i_1, \dots, i_{n(\varepsilon)}} \sigma^{\varepsilon}(i_1, \dots, i_{n(\varepsilon)}) \\ & \times \ln \sup_z P^{\varepsilon}_z\{x^{\varepsilon}_1 \in V_{i_1}, \dots, x_{n(\varepsilon)} \in V_{i_{n(\varepsilon)}}\} \\ \geqslant & - \sum_{(i_1, \dots, i_{n(\varepsilon)}) \in G_{n(\varepsilon)}} \sigma^{\varepsilon}(i_1, \dots, i_{n(\varepsilon)}) \\ & \times \ln \sup_z P^{\varepsilon}_z\{x^{\varepsilon}_1 \in V_{i_1}, \dots, x^{\varepsilon}_{n(\varepsilon)} \in V_{i_{n(\varepsilon)}}\}, \end{aligned} \tag{4.30}$$

where $G_{n(\varepsilon)}$ is the collection of $n(\varepsilon)$-sequences $(i_1, \dots, i_{n(\varepsilon)})$ such that there exists an integer $r = r_{i_1, \dots, i_{n(\varepsilon)}} \leqslant \nu$ so that $\Lambda \subset V_{i_j} \subset U_{2\delta}(\Lambda_r) \subset U_{\Lambda_r}$ for all $j = 1, \dots, n(\varepsilon)$, where $\delta > \rho > 0$ is small enough.

Then clearly,

$$\sum_{(i_1, \dots, i_{n(\varepsilon)}) \notin G_{n(\varepsilon)}} \sigma^{\varepsilon}(i_1, \dots, i_{n(\varepsilon)}) \leqslant n(\varepsilon) \mu^{\varepsilon}_B\Big(M \setminus \bigcup_{1 \leqslant j \leqslant \nu} U_{\delta}(\Lambda_j)\Big). \tag{4.31}$$

It follows from [**12**] (see §8 and Remark 3 there) that for some $C^{(3)}$, $\alpha_3 > 0$,

$$\mu^{\varepsilon}_B\Big(M \setminus \bigcup_{1 \leqslant j \leqslant \nu} U_{\delta}(\Lambda_j)\Big) \leqslant C^{(3)} \exp\Big(\frac{-\alpha_3}{\varepsilon^2}\Big). \tag{4.32}$$

Therefore,

$$\sum_{(i_1, \dots, i_{n(\varepsilon)}) \in G_{n(\varepsilon)}} \sigma^{\varepsilon}(i_1, \dots, i_{n(\varepsilon)}) \to 1 \quad \text{as } \varepsilon \to 0, \tag{4.33}$$

so one can consider the process x^{ε}_t, $0 \leqslant t \leqslant n(\varepsilon)$, separately in each neighborhood $U_{2\delta}(\Lambda_j), j = 1, \dots, \nu$. Then we apply the proof of Theorem 1 to show that if

$$\mu^{\varepsilon_i}_B\big(U_{2\delta}(\Lambda_j)\big) \to q_j \quad \text{as } \varepsilon_i \to 0 \tag{4.34}$$

then

$$\frac{1}{q_j} \lim_{\varepsilon_i \to 0} \frac{1}{n(\varepsilon_i)} \tilde{H}^{\varepsilon}_{n(\varepsilon_i)}(\Gamma_{\Lambda_j}) \geqslant h_{\mu_{\Lambda}}(S^1). \tag{4.35}$$

Therefore using (1.12), (2.11), (2.12), (4.16) and (4.28), one gets (2.15).

5. Concluding remarks. It follows from [**3**] that for m-almost all point y from the neighborhood U_{Λ} of a hyperbolic attractor Λ,

$$\lim_{t \to \infty} \frac{1}{t} \ln J_t(y) = h_{\mu_{\Lambda}}(S^1). \tag{5.1}$$

Set

$$Q_x^\varepsilon(T,\delta) = P_x^\varepsilon\Big\{\sup_{0\le t\le T} \operatorname{dist}\big(x_t^\varepsilon, S^t x\big) \le \delta\Big\}.$$

Then by (5.1) and Theorem 2.1 of [**9**] for m-almost all $x \in U_\delta(\Lambda)$,

$$\lim_{\varepsilon\to 0}\limsup_{T\to\infty}\frac{1}{T}\ln Q_x^\varepsilon(T,\delta) = -h_{\mu_\Lambda}(S^1) \tag{5.2}$$

provided $\delta > 0$ is small enough. Hence

$$h_{\mu_\Lambda}(S^1) = -\frac{1}{m(U_\delta(\Lambda))}\int_{U_\delta(\Lambda)}\lim_{\varepsilon\to 0}\limsup_{T\to\infty}\frac{1}{T}(\ln Q_x^\varepsilon(T,\delta))m(dx). \tag{5.3}$$

If M is compact and L^ε is a nondegenerate, then the distribution $P_y^\varepsilon\{x_1^\varepsilon \in dx\}$ is equivalent to the Riemannian volume, so one can replace $m(dx)$ in (5.3) by this distribution. The shortcoming of this formula is its use of the trajectory of the flow S^t for the representation of the entropy. On the contrary the left-hand sides of (2.9) and (2.15) do not depend on the flow S^t but only on its random perturbations x_t^ε.

Some version of this can be proved for the nonuniform hyperbolic case, namely for flows with smooth invariant measures and nonzero characteristic exponents (see [**11**]). Then there is a function $\rho(x)$ positive on regular points such that in neighborhoods $U_{\rho(x)}(x)$ the dynamical system has properties similar to the uniform hyperbolic case (see [**7**] in the case of diffeomorphisms) and this function satisfies (see [**10**])

$$\int_M |\ln \rho(x)|\mu(dx) < \infty, \tag{5.4}$$

where μ is a smooth S^t-invariant erogodic measure.

From (5.4) it follows that the times t for which $\rho(S^t x) < \varepsilon^\beta$, $0 < \beta < 1$, are rare. Hence, proceeding in the same way as in [**9**], one can show that if

$$Q_x(T,\rho,\delta) = P_x^\varepsilon\big\{\operatorname{dist}\big(x_t^\varepsilon, S^t x\big) \le \min\big(\delta, \rho(S^t x)\big)\forall\, t \in [0,T]\big\}$$

and x is a regular point, then

$$-\lim_{\varepsilon\to 0}\limsup_{T\to\infty}\frac{1}{T}\ln Q_x^\varepsilon(T,\rho,\delta) = \lambda^+(x) = h_\nu(S^1), \tag{5.5}$$

where $\lambda^+(x)$ is the sum of the positive characteristic exponents at x (see [**11**]). Since the distribution of x_1^ε is equivalent to the Riemannian volume m, provided M is compact and L^ε is nondegenerate, then one has for any $y \in M$,

$$\begin{aligned} h_\mu(S^1) &= -\lim_{\varepsilon\to 0} E_y^\varepsilon \limsup_{T\to\infty}\frac{1}{T}\ln Q_{x_1^\varepsilon}(T,\rho,\delta) \\ &= -\lim_{\varepsilon\to 0}\int_M \limsup_{T\to\infty}\frac{1}{T}\ln Q_x^\varepsilon(T,\rho,\delta)m(dx), \end{aligned} \tag{5.6}$$

where E_x^ε is the expectation for the process x_t^ε starting at x.

REFERENCES

1. D. G. Aronson, *Non-negative solutions of linear parabolic equations*, Ann. Scuola Norm. Sup. Pisa (3) **22** (1968), 607–697.

2. P. Billingsley, *Ergodic theory and information*, Wiley, New York, 1965.

3. R. Bowen and D. Ruella, *The ergodic theory of axiom* A *flows*, Invent. Math. **29** (1975), 181–202.

4. J. Doob, *Stochastic processes*, Wiley, New York, 1953.

5. R. Z. Khas'minskiĭ, *The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion*, Theor. Probability Appl. **8** (1963), 1–21.

6. N. Ikeda and A. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland, Amsterdam, 1981.

7. A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math. **51** (1980), 137–173.

8. Yu. Kifer, *On small random perturbations of some smooth dynamical systems*, Math. USSR-Izv. **8** (1974), 1083–1107.

9. _____, *Stochastic stability of the topological pressure*, J. Analyse Math. **38** (1980), 255–286.

10. R. Mañé, *A proof of Pesin's formula*, Ergodic Theory Dynamical Systems **1** (1981), 95–102.

11. Ja. B. Pesin, *Characteristic Lyapunov exponents and smooth ergodic theory*, Russian Math. Surveys **32** (1977), 55–114.

12. A. D. Ventcel and M. I. Freidlin, *On small random perturbations of dynamical systems*, Russian Math. Surveys **25** (1970), 1–56.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742

INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM, ISRAEL (Current address)