ENTROPY VIA RANDOM PERTURBATIONS

BY

YURI KIFER

Abstract. The entropy of a dynamical system $S'$ on a hyperbolic attractor with respect to the Bowen-Ruelle-Sinai measure is obtained as a limit of entropy characteristics of small random perturbations $x^e$ of $S'$. 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 $S': M \to M$, $-\infty < t < \infty$, defined by

$$d(S'x)/dt = B(S'x),$$

and the family of diffusion Markov processes $x^e$ given by the stochastic differential equation

$$dx^e_t = B(x^e_t) dt + \epsilon \sum_{1 \leq i \leq d} X_i(x^e_t) dw^i_t,$$

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

$$L^e = B + \frac{\epsilon^2}{2} \sum_{1 \leq i \leq d} X_i X_i.$$

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

$$f(x^e_t) = f(x^0_t) + \int_0^t Bf(x^e_s) ds + \epsilon \sum_{1 \leq i \leq d} \int_0^t X_i f(x^e_s) dw^i_s.$$

The processes $x^e_t$ are called small random perturbations of the flow $S'$. 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 $x^e_t$ 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.

Supported by U.S.A.–Israel BSF.
Let $\Gamma = (V_1, \ldots, V_k)$ be a partition of $M$ into closed subsets such that
\begin{equation}
\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,
\end{equation}
where $\partial V_i = V_i \setminus \text{int} V_i$ is the boundary of $V_i$. Denote
\begin{equation}
P^t(x, Q) = P^t_x\{x' \in Q\},
\end{equation}
for any Borel subset $Q$, where $P^t_x\{\cdot\}$ is the probability of the event in brackets, provided $x'_0 = x$. Suppose $x'_t$ has an invariant probability measure $\mu^t$, i.e.
\begin{equation}
\mu^t(Q) = \int_M \mu^t(dx) P^t(x, Q)
\end{equation}
for any Borel subset $Q$ and $t > 0$. Then one can consider
\begin{equation}
\sigma^t(i_0, \ldots, i_n) = \int_{V_{i_0}} \mu^t(dx) P^t_x\{x'_1 \in V_{i_1}, x'_2 \in V_{i_2}, \ldots, x'_n \in V_{i_n}\}.
\end{equation}
Assume
\begin{equation}
\mu^t(\partial \Gamma) = 0,
\end{equation}
where $\partial \Gamma = \bigcup_{i=1}^{k} \partial V_i$ and set
\begin{equation}
H^t_n(\Gamma) = - \sum_{i_0, \ldots, i_n} \sigma^t(i_0, \ldots, i_n) \ln \sigma^t(i_0, \ldots, i_n),
\end{equation}
where $0 \ln 0 \equiv 0$. It is known (see [2]) that
\begin{equation}
H^t_n(\Gamma) + H^t_n(\Gamma) = H^t_{n+n}(\Gamma),
\end{equation}
so by the standard subadditivity argument the limit
\begin{equation}
H^t(\Gamma) = \lim_{N \to \infty} \frac{1}{N + 1} H^t_N(\Gamma) = \inf_{N \geq 0} \frac{1}{N + 1} H^t_N(\Gamma)
\end{equation}
eexists. The quantity $H^t(\Gamma)$ is the entropy (see [2]) of the stationary Markov chain $x^t_n$, $n = 0, 1, \ldots$, with the invariant measure $\mu^t$ with respect to the partition of the sample space $\Omega$ into the sets
\begin{equation}
\Omega_i = \{\omega: x^t_0(\omega) \in V_i\}, \quad i = 1, \ldots, k.
\end{equation}
Assume
\begin{equation}
\mu^t \to \mu \quad \text{as} \ \epsilon_j \to 0 \ \text{in the weak sense} \quad \left(\mu^t \overset{w}{\to} \mu\right)
\end{equation}
and
\begin{equation}
\limsup_{\epsilon \to 0} \int_M |P^t(t, x, dy)f(y) - f(S'tx)| = 0
\end{equation}
for any $t > 0$ and a continuous function $f$ on $M$. Then (see [5]) the measure $\mu$ will be invariant for the flow $S'$. If
\begin{equation}
\mu(\partial \Gamma) = 0
\end{equation}
then one can also consider the entropy $H(\Gamma)$ of the transformation $S^1$ with respect to the partition $\Gamma$ (see [2]),

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

where

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

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

$$\sigma(i_0, \ldots, i_N) = \mu \left( \bigcap_{0 \leq j \leq N} S^{-j} V_i \right).$$

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

$$\sigma^{\epsilon_j}(i_0, \ldots, i_N) \to \sigma(i_0, \ldots, i_N) \quad \text{as} \ \epsilon_j \to 0,$$

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

$$\limsup_{\epsilon_j \to 0} H_{\epsilon_j}(\Gamma) \leq \inf_{N \geq 0} \frac{1}{N + 1} \limsup_{\epsilon_j \to 0} H_N(\Gamma) = H(\Gamma).$$

We shall prove in this paper that if $S'$ is a hyperbolic flow, $\Gamma$ is a partition into elements of small size satisfying (1.9) and (1.14), and $\mu$ is the Sinai-Bowen-Ruelle measure on an attractor, then

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

where $h_{\mu}(S^1) = \sup_{\Gamma} H(\Gamma)$ is the entropy of $S^1$ with respect to the measure $\mu$ (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 $\Lambda$, i.e., when the operator $L'$ coincides with $B$ outside of some neighborhood of $\Lambda$ and $L'$ is a nondegenerate elliptic operator in some smaller neighborhood of $\Lambda$. 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, $x_i'$ is a nondegenerate diffusion on $M$, and the $\omega$-limit set of $S'$ is the disjoint union of a finite number of hyperbolic attractors and a finite number of closed unstable sets. The measure $\mu$ in this case is a combination of Sinai-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 $P'(t, x, Q)$, such as Döblin’s condition (see [4]), which are satisfied provided the
operator $L^t$ is nondegenerate and $M$ is compact, the entropy of the Markov chain $x_n$, i.e. $\sup T \mathcal{H}^t(\Gamma)$, 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 $\epsilon \to 0$ 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 $\Lambda$ can be written as the Whitney sum of continuous subbundles

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

where $F^0 = B$, this decomposition is invariant with respect to the differential $DS'$ of $S^t$, and there exist $c, \nu > 0$ such that

$$\|DS'^t \xi\| \leq c e^{-\nu t} \|\xi\| \quad \text{for} \ \xi \in F^s, \ t \geq 0,$$

and

$$\|DS'^{-t} \eta\| \leq c e^{\nu t} \|\eta\| \quad \text{for} \ \eta \in F^u, \ t \geq 0.$$

A set $\Lambda$ is called a hyperbolic (basic) attractor if:

(a) $\Lambda$ is hyperbolic and either $\Lambda$ is a point or it contains no fixed points;
(b) the periodic orbits of $S'|_{\Lambda}$ are dense in $\Lambda$;
(c) $S'|_{\Lambda}$ is a topologically transitive flow;
(d) there is an open set $U_{\Lambda} \supset \Lambda$ such that

$$S'^t U_{\Lambda} \subset U_{\Lambda} \quad \text{for all} \ t > 0 \quad \text{and} \quad \Lambda = \bigcap_{t \geq 0} S'^t U_{\Lambda}.$$

For $x \in \Lambda$ let $J_t(x)$ be the Jacobian of the linear map $DS'^t: F^u_x \to F^u_{S'^t x}$. Define

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

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

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

**Assumption A.** Let $\Lambda$ be a hyperbolic attractor, an open set $U_{\Lambda}$ has a compact closure and satisfies (2.4), the operator $L^t$ defined in (1.3) is a nondegenerate elliptic operator in some open set $U_{\Lambda} \subset U_{\Lambda}$ and $L^t \equiv B |_{M \setminus U_{\Lambda}}$. 
If Assumption A is satisfied then it follows that the process $x_t^\varepsilon$ never leaves $U_\Lambda \equiv U_\Lambda \cup \partial U_\Lambda$ provided $x_0^\varepsilon \in U_\Lambda$. Using [8] we shall prove the following result in §3.

**Proposition 1.** Suppose Assumption A is valid. Then the process $x_t^\varepsilon$ has the unique invariant measure $\mu_A^\varepsilon$ having support in $U_\Lambda$ and

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

The main result of this paper is the following

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

$$\partial \Gamma \subset U_L, \quad m(\partial \Gamma) = \mu_\Lambda(\partial \Gamma) = 0$$  

and $\max_{V_i \in \Gamma_\Lambda(\rho)} \text{diam} V_i \leq \rho$, where $\Gamma_\Lambda(\rho) \equiv \{V_i: V_i \cap U_\rho(\Lambda) \neq \emptyset\}$, $U_\rho(\Lambda) \equiv \{y: \text{dist}(y, \Lambda) \leq \rho\}$, $\rho > 0$ is small enough and $m$ is the Riemannian volume.

Then

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

where $H^r(\Gamma)$ 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 $\Gamma$ 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 $S^r$ but only on the process $x_t^\varepsilon$.

For any $x, y \in M$ define

$$d(x, y) = \inf_{0 \leq \tau \leq 1} \inf_{x_0 = S^\tau x} \sum_{0 \leq i \leq n-1} \text{dist}(S^\tau x_i, x_{i+1})^2,$$

where $\inf$ is taken over all finite sequences $\{x_i \in M, 1 \geq \tau \geq 0, i = 0, \ldots, n; x_0 = S^\tau x$ 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 $L^* \in \mathcal{D}$ defined by (1.3) is a nondegenerate elliptic operator on $M$; (b) there exist a finite number of equivalence classes $K_1, \ldots, K_\ell$ (with respect to the relation $\sim$ above) which are compact sets and contain any $\omega$-limit set of the equation (1.1); (c) the stable sets among $K_1, \ldots, K_\ell$ are $K_i = \Lambda_1, \ldots, K_\ell = \Lambda_\ell$ 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 $P^*(t, x, Q)$ of the process $x^t$ 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 $\mu_B$. It follows from [8 and 12] under conditions of Assumption B that if

$$\mu_B^\varepsilon \xrightarrow{w} \mu_0 \quad \text{as } \varepsilon \to 0,$$

then

$$\mu_0 = \sum_{1 \leq j \leq \rho} q_j \mu_{\Lambda_j},$$
where \( q_j \geq 0, \Sigma_{1 \leq j \leq r} q_j = 1 \) 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 \( \Gamma = (V_1, \ldots, V_k) \) be a partition of \( M \) such that

\[
m(\partial \Gamma) = \mu_{\Lambda_i}(\partial \Gamma) = 0 \quad \text{for all } i = 1, \ldots, r
\]

and

\[
\max_i \max_{V_j \in \Gamma_i} \text{diam } V_j \leq \rho,
\]

where \( \rho > 0 \) is small enough and \( \Gamma_i \) is defined in the same way as in Theorem 1. Then

\[
\lim_{t_i \to 0} H^{t_i}(\Gamma) = \sum_{1 \leq j \leq r} h_{\mu_{\Lambda_j}}(S^t),
\]

where \( \{ \epsilon_i \} \) is the same sequence as in (2.11) and \( H^t(\Gamma) \) is defined by (1.8), (1.10) and (1.12) with \( \mu^t = \mu_B^t \).

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 \( \partial U_2 \) such that

\[
\Lambda \subset U_0 \subset U_0 \cup \partial U_0 \subset U_1 \subset U_1 \cup \partial U_1 \subset U_2 \subset U_2 \cup \partial U_2 \subset U_L.
\]

Suppose that Assumption A is satisfied. Then \( x^t \in U_\Lambda \cup \partial U_\Lambda \) for all \( t \geq 0 \) provided \( x_0^t \in U_\Lambda \cup \partial U_\Lambda \).

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

\[
q^t = \inf_{x \in U_\Lambda} P^t(T, x, U_1) \to 1 \quad \text{as } \epsilon \to 0.
\]

Let \( \tau_0 = 0, \) let \( \tau_1 \) be the first time the process \( x^t \) hits \( \partial U_1 \), and let \( \tilde{\tau}_1 \) be the first time after \( \tau_1 \) that \( x^t \) hits \( \partial U_2 \). Let \( \tau_i(\tilde{\tau}_i) \) be the first time after \( \tilde{\tau}_{i-1}(\tau_{i-1}) \) that \( x^t \) hits \( \partial U_1(\partial U_2) \).

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

\[
P^t(t, x, V) = \tilde{P}^t(t, x, V) + E_{x}^{\tau_{\tau_{1-1}}} P^t(t - \tau_2, x^t_{\tau_2}, V)
\]

\[
= \sum_{n=0}^{\infty} E_{x}^{\tau_{\tau_{1-1}}} \tilde{P}^t(t - \tau_n, x^t_{\tau_n}, V),
\]

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

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

\[
\tilde{q}^t \geq \tilde{p}^t(1, x, y) \geq \tilde{q}^t \quad \text{for any } x, y \in U_1 \cup \partial U_1
\]
and
\[(3.5) \quad \tilde{p}^t(t, x, y) \leq \tilde{q}^t \quad \text{for any } t \geq 0 \text{ if } x \in \partial U_1 \text{ and } y \in U_0,\]
where \(\tilde{p}^t(t, x, y)\) is the transition density of the process \(x'_t\) with absorption on \(\partial U_2\), i.e.
\[(3.6) \quad \tilde{P}^t(t, x, V) = \int_V \tilde{p}^t(t, x, y) m(\text{d}y)\]
for any Borel set \(V\).

From the strong Markov property and (3.2) it follows that
\[(3.7) \quad \sup_{x \in U_1} P_x^t(\tau_n < T) \leq (1 - q')^n.\]
Therefore by (3.3)–(3.7),
\[(3.8) \quad (1 + 1/q') q' m(V) \geq P^t(T + 1, x, V) \geq q' \cdot q^t m(V).\]

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

Since by the Chapman-Kolmogorov formula,
\[(3.9) \quad p^t(t, x, V) = \int_{U_1} p^t(t - T - 1, x, dy) P^t(T + 1, y, V),\]
then \(P^t(t, x, V)\) has a positive density \(p^t(t, x, y)\) for any \(t \geq T + 1\).

Therefore the process \(x'_t\) satisfies the Döblin condition (see [4, Chapter 5]), so it has a unique invariant measure \(\mu^t_\Lambda\) with support in \(\bar{U}_1\). Moreover, \(\mu^t_\Lambda\) has a positive density \(g^t_\Lambda(x)\) in \(U_1\) with respect to the Riemannian volume \(m\) and this density is bounded away from zero and infinity in any compact subset of \(U_1\). Hence, under the conditions of Theorem 1, when \(m(\partial \Gamma) = 0\) and \(\partial \Gamma \subset U_1\), one obtains
\[(3.10) \quad \mu^t_\Lambda(\partial \Gamma) = 0.\]

The reduction of the transition probabilities \(P^t(t, x, V)\) in (3.3) to transition probabilities of the process \(x'_t\) with absorption on \(\partial U_2\) enables us to consider \(x'_t\) 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 \(x'_t\) near \(\Lambda\). This remark completes the proof of Proposition 1.

4. Proofs of Theorems 1 and 2. Let
\[(4.1) \quad d_T(x, y) = \sup_{0 \leq t \leq T} \text{dist}(S'tx, S'ty)\]
and
\[(4.2) \quad B_x(\delta, T) = \{ y \in M' : d_T(x, y) \leq \delta \}.\]

According to [3] a subset \(E \subset \Lambda\) is called \((\delta, T)\)-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 \((\delta, T)\)-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 \( \delta \) and \( T \). Thus there exists a maximal \((\delta, T)\)-separated set \( E(\delta, T) \) (not necessarily unique).

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

\[
\bigcup_{x \in E(\delta, T)} B_\delta(2\delta, T) \cup \bigcup_{x \in \Lambda} B_\delta(\delta, T) \supset U_{\alpha_1, \delta}(\Lambda) \supset \Lambda,
\]

where \( \alpha_1 > 0 \) is independent of \( \delta, T \).

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

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

Define

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

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

\[
m(B_\delta(\delta, T)) \leq C_\delta \mu_\Lambda(B_\delta(\delta, T)),
\]

\[
C_\delta^{-1} \leq m(B_\delta(\delta, T)) J_T(x) \leq C_\delta,
\]

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

\[
C_\delta^{-1} \leq J_T(x) J_T^{-1}(y) \leq C_\delta.
\]

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

\[
T > (1/\gamma) \ln C_\delta,
\]

then by (4.8),

\[
B_\delta(\delta, T) \cap \Lambda \subset Q_\gamma(T).
\]

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

\[
\mu_\Lambda(Q_\gamma(T)) \geq \mu_\Lambda \left( \bigcup_{x \in E(\delta, T) \cap Q_{2\gamma}(T)} B_\delta(\delta, T) \right)
\]

\[
\geq C_\delta^{-2} \sum_{x \in E(\delta, T) \cap Q_{2\gamma}(T)} J_T^{-1}(x).
\]

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

\[
\max_{V_i \in \Gamma_\rho} \text{diam} V_i \leq \rho \leq \frac{1}{3} \alpha_1 \delta \leq \frac{1}{3} \delta.
\]

Define \( U_\rho^\Gamma = \bigcup_{V_i \in \Gamma_\rho} V_i \). Then

\[
\Lambda \subset U_\rho^\Gamma \subset U_{\alpha_1, \delta}(\Lambda).
\]

Choose \( \beta > 0 \) small enough and set

\[
n(\epsilon) = \text{integral part of } \epsilon^{-\beta}.
\]
By the Markov property
\begin{equation}
\sigma(i_0, \ldots, i_{N\cdot n(\varepsilon)}) \leq \sup_{z} P^t_z\{ x^t_1 \in V_{i_1}, \ldots, x^t_{n(\varepsilon)} \in V_{i_{n(\varepsilon)}} \} \times \sup_{z} P^t_z\{ x^t_1 \in V_{i_{n(\varepsilon)}+1}, \ldots, x^t_{2n(\varepsilon)} \in V_{2n(\varepsilon)} \} \times \cdots \sup_{z} P^t_z\{ x^t_1 \in V_{i_{N\cdot n(\varepsilon)}+1}, \ldots, x^t_{N\cdot n(\varepsilon)} \in V_{N\cdot n(\varepsilon)} \},
\end{equation}
where the sup is taken over \( z \in \overset{\_}{U}_{\Lambda} \). Therefore
\begin{equation}
H_{N\cdot n(\varepsilon)}^{\varepsilon}(\Gamma) = - \sum_{i_0, \ldots, i_{N\cdot n(\varepsilon)}} \sigma(i_0, \ldots, i_{N\cdot n(\varepsilon)}) \ln \sigma(i_0, \ldots, i_{N\cdot n(\varepsilon)})
\end{equation}
\begin{equation}
\geq - \sum_{k=1}^{N} \left( \sum_{i_{k-1\cdot n(\varepsilon)+1}, \ldots, i_{k\cdot n(\varepsilon)}} \sigma(i_{k-1\cdot n(\varepsilon)+1}, \ldots, i_{k\cdot n(\varepsilon)}) \right)
\times \ln \sup_{z} P^t_z\{ x^t_1 \in V_{i_{k-1\cdot n(\varepsilon)+1}}, \ldots, x^t_{n(\varepsilon)} \in V_{i_{k\cdot n(\varepsilon)}} \}
= -N \sum_{i_1, \ldots, i_{n(\varepsilon)}} \sigma(i_1, \ldots, i_{n(\varepsilon)}) \ln \sup_{z} P^t_z\{ x^t_1 \in V_{i_1}, \ldots, x^t_{n(\varepsilon)} \in V_{i_{n(\varepsilon)}} \}
\end{equation}
def \( \equiv N H_{n(\varepsilon)}^{\varepsilon}(\Gamma) \).

Moreover,
\begin{equation}
\tilde{H}_{n(\varepsilon)}^{\varepsilon}(\Gamma) \geq \mathcal{H}^{\varepsilon}(\mathcal{G}_{n(\varepsilon)})
\end{equation}
def \( \equiv \sum_{(i_1, \ldots, i_{n(\varepsilon)}) \in \mathcal{G}_{n(\varepsilon)}} \sigma(i_1, \ldots, i_{n(\varepsilon)}) \ln \sup_{z} P^t_z\{ x^t_1 \in V_{i_1}, \ldots, x^t_{n(\varepsilon)} \in V_{i_{n(\varepsilon)}} \},
\end{equation}
where \( \mathcal{G}_{n(\varepsilon)} \) is the collection of \( n(\varepsilon) \)-sequences \( i_1, \ldots, i_{n(\varepsilon)} \) such that if \( (i_1, \ldots, i_{n(\varepsilon)}) \in \mathcal{G}_{n(\varepsilon)} \), then
(i) \( V_{i_1} \cap U_{\delta}(\Lambda) \neq \emptyset \), i.e. \( V_{i_1} \in \Gamma_{\Lambda}(\rho) \) for all \( j = 1, \ldots, n(\varepsilon) \);
(ii) there exists \( x \in E(\delta, n(\varepsilon)) \) such that \( V_{i_j} \subset U_{2\delta}(S^jx), j = 1, \ldots, n(\varepsilon) \), where \( E(\delta, n(\varepsilon)) \) is a maximal \( (\delta, n(\varepsilon)) \)-separated set.
(iii) the point \( x \) from (ii) does not belong to \( Q_{2\delta}(n(\varepsilon)) \).

It follows from [8, 9 and 12] that
\begin{equation}
\mu(\overset{\_}{U}_{\Lambda} \setminus U^{\Gamma}_{\Lambda}) \leq C^{(1)} \exp(-\alpha_2/\varepsilon^2),
\end{equation}
and for any \( z \in \overset{\_}{U}_{\Lambda} \),
\begin{equation}
P^t_z\left\{ \inf_{y \in U_{\delta}(z)} \sup_{0 < r < n(\varepsilon)} \text{dist}(x^t_r, S^r y) \geq \rho \right\} \leq C^{(1)} \exp\left( -\frac{\alpha_2}{\varepsilon} \right)
\end{equation}
for some \( C^{(1)}, \alpha_2 > 0 \) independent of \( \varepsilon \), provided \( \beta \) in (4.14) is small enough.

On the other hand, if \( \delta \) 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
\begin{equation}
P^t_{y}\left\{ \max_{1 \leq j < n(\varepsilon)} \text{dist}(x^t_j, S^j x) \leq 2\delta \right\} \leq C^{(2)} \varepsilon^{-d} J_{n(\varepsilon)}^{-1}(x)
\end{equation}
and
\[ (4.21) \quad \int_{U_\lambda} \mu^\epsilon(dy) P_{\gamma}^\epsilon \left\{ \max_{1 \leq j \leq n(\epsilon)} \text{dist}(x_j, S'x) \leq 2\delta \right\} \leq C^{(2)} f_{n(\epsilon)}(x) \]

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

From (4.20), together with (i)–(iii), it follows that
\[ (4.22) \quad \frac{1}{n(\epsilon)} \beta^\epsilon(S_{\gamma, n(\epsilon)}) \geq \left( h_{\mu^\epsilon}(S^1) + \frac{d_u}{n(\epsilon)} \ln \epsilon - 2\gamma - \frac{1}{n(\epsilon)} \ln C^{(2)} \right) \times \sum_{(i_1, \ldots, i_{n(\epsilon)}) \in S_{\gamma, n(\epsilon)}} \sigma^\epsilon(i_1, \ldots, i_{n(\epsilon)}). \]

Next from (4.3), (4.11), (4.12), (4.18), (4.19) and (4.21), one obtains
\[ (4.23) \quad \sum_{(i_1, \ldots, i_{n(\epsilon)}) \in S_{\gamma, n(\epsilon)}} \sigma^\epsilon(i_1, \ldots, i_{n(\epsilon)}) \leq n(\epsilon) \mu^\epsilon(U_{\lambda} \setminus U_{\Gamma}) + \sup \left\{ \int_{U_\lambda} \mu^\epsilon(dy) P_{\gamma}^\epsilon \left\{ \max_{1 \leq j \leq n(\epsilon)} \text{dist}(x_j, S'x) \leq 2\delta \right\} \right\} \]
\[ + \sum_{x \in E(\delta, \Gamma) \cap Q_{\gamma, n(\epsilon)}} \int_{U_\lambda} \mu^\epsilon(dy) P_{\gamma}^\epsilon \left\{ \text{dist}(x_j, S'x) \leq 2\delta \right\} \]
\[ \leq (n(\epsilon) + 1) \exp(-\alpha_2/\epsilon) + C^{(2)} C_{\delta, 2}^2 \mu^\epsilon(Q_{\gamma}(n(\epsilon))). \]

From (4.4) it follows that
\[ (4.24) \quad \lim_{\epsilon \to 0} \mu_{\lambda}^\epsilon(Q_{\gamma}(n(\epsilon)) = 0. \]

By (3.10),
\[ \sum_{i_1, \ldots, i_{n(\epsilon)}} \sigma^\epsilon(i_1, \ldots, i_{n(\epsilon)}) = 1, \]
so (4.22)–(4.24) yield
\[ (4.25) \quad \liminf_{\epsilon \to 0} \frac{1}{n(\epsilon)} \beta^\epsilon(S_{\gamma, n(\epsilon)}) \geq h_{\mu^\epsilon}(S^1) - 2\gamma. \]

Since $\gamma$ can be arbitrarily small, one obtains from here and (4.16)–(4.17) that
\[ (4.26) \quad \liminf_{\epsilon \to 0} \frac{1}{n(\epsilon)} \hat{H}_{\epsilon, n(\epsilon)}(\Gamma) \geq h_{\mu^\epsilon}(S^1). \]

On the other hand, the limit (1.12) exists, so by (4.16),
\[ (4.27) \quad H^\epsilon(\Gamma) \geq (1/n(\epsilon)) \hat{H}_{\epsilon, n(\epsilon)}(\Gamma), \]
and by (4.26),
\[ (4.28) \quad \liminf_{\epsilon \to 0} H^\epsilon(\Gamma) \geq h_{\mu^\epsilon}(S^1). \]

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 \( \Gamma = \{V_1, \ldots, V_k\} \) is a corresponding partition of \( M \). Then by (1.19), (2.11) and (2.12),

\[
(4.29) \quad \lim_{\epsilon \to 0} \sup \, H^\epsilon (\Gamma) \leq H(\Gamma) \leq h_{\mu_\Lambda}(S^1) = \sum_{1 \leq j \leq \nu} q_j h_{\mu_\Lambda}(S^1).
\]

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).

(4.30) \[
\hat{H}_{n(\epsilon)}^\epsilon (\Gamma) \equiv - \sum_{(i_1, \ldots, i_{n(\epsilon)})} \sigma^\epsilon (i_1, \ldots, i_{n(\epsilon)}) \times \ln \sup_z P_z^\epsilon \{x_i^\epsilon \in V_i, \ldots, x_{n(\epsilon)}^\epsilon \in V_{n(\epsilon)}\} \\
\geq - \sum_{(i_1, \ldots, i_{n(\epsilon)}) \in G_{n(\epsilon)}} \sigma^\epsilon (i_1, \ldots, i_{n(\epsilon)}) \times \ln \sup_z P_z^\epsilon \{x_i^\epsilon \in V_i, \ldots, x_{n(\epsilon)}^\epsilon \in V_{n(\epsilon)}\},
\]

where \( G_{n(\epsilon)} \) is the collection of \( n(\epsilon) \)-sequences \((i_1, \ldots, i_{n(\epsilon)})\) such that there exists an integer \( r = r_{i_1, \ldots, i_{n(\epsilon)}}, \leq \nu \) so that \( \Lambda \subset V_j \subset U_{2\delta}(\Lambda_j) \subset U_{\Lambda_j} \) for all \( j = 1, \ldots, n(\epsilon) \), where \( \delta > \rho > 0 \) is small enough.

Then clearly,

(4.31) \[
\sum_{(i_1, \ldots, i_{n(\epsilon)}) \not\in G_{n(\epsilon)}} \sigma^\epsilon (i_1, \ldots, i_{n(\epsilon)}) \leq n(\epsilon) \mu_\beta^\epsilon \left( M \setminus \bigcup_{1 \leq j \leq \nu} U_{\delta}(\Lambda_j) \right).
\]

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

(4.32) \[
\mu_\beta^\epsilon \left( M \setminus \bigcup_{1 \leq j \leq \nu} U_{\delta}(\Lambda_j) \right) \leq C(3) \exp \left( -\frac{\alpha_3}{\epsilon^2} \right).
\]

Therefore,

(4.33) \[
\sum_{(i_1, \ldots, i_{n(\epsilon)}) \in G_{n(\epsilon)}} \sigma^\epsilon (i_1, \ldots, i_{n(\epsilon)}) \to 1 \quad \text{as} \quad \epsilon \to 0,
\]

so one can consider the process \( x_i^\epsilon, 0 \leq i \leq n(\epsilon) \), separately in each neighborhood \( U_{2\delta}(\Lambda_j), j = 1, \ldots, \nu \). Then we apply the proof of Theorem 1 to show that if

(4.34) \[
\mu_\beta^\epsilon (U_{2\delta}(\Lambda_j)) \to q_j \quad \text{as} \quad \epsilon \to 0
\]

then

(4.35) \[
\frac{1}{q_j} \lim_{\epsilon \to 0} \frac{1}{n(\epsilon)} \hat{H}_{n(\epsilon)}^\epsilon (\Gamma_{\Lambda_j}) \geq h_{\mu_\Lambda}(S^1).
\]

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 \( U_{\Lambda} \) of a hyperbolic attractor \( \Lambda \),

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

\[ Q^\varepsilon(T, \delta) = P^\varepsilon\left\{ \sup_{0 \leq t \leq T} \text{dist}(x^\varepsilon_t, S'tx) \leq \delta \right\}. \]

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

\[
\lim_{\varepsilon \to 0} \limsup_{T \to \infty} \frac{1}{T} \ln Q^\varepsilon_x(T, \delta) = -h_{\mu}(S')
\]

provided \( \delta > 0 \) is small enough. Hence

\[
h_{\mu}(S') = \frac{1}{m(U_\delta(\Lambda))} \int_{U_\delta(\Lambda)} \lim_{\varepsilon \to 0} \limsup_{T \to \infty} \frac{1}{T} (\ln Q^\varepsilon_x(T, \delta) - m(dx)).
\]

If \( M \) is compact and \( L' \) is a nondegenerate, then the distribution \( P_x^\varepsilon(\{ x^\varepsilon_t \in dx \}) \) 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 \( S' \) 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 \( S' \) but only on its random perturbations \( x^\varepsilon_t \).

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])

\[
\int_{M} |\ln \rho(x)| \mu(dx) < \infty,
\]

where \( \mu \) is a smooth \( S' \)-invariant ergodic measure.

From (5.4) it follows that the times \( t \) for which \( \rho(S'tx) < e^{\beta} \), \( 0 \leq \beta < 1 \), are rare. Hence, proceeding in the same way as in [9], one can show that if

\[
Q_x(T, \rho, \delta) = P_x^\varepsilon\left( \{ \text{dist}(x^\varepsilon_t, S'tx) \leq \min(\delta, \rho(S'tx)) \forall t \in [0, T] \} \right)
\]

and \( x \) is a regular point, then

\[
-\lim_{\varepsilon \to 0} \limsup_{T \to \infty} \frac{1}{T} \ln Q_x^\varepsilon(T, \rho, \delta) = \lambda^+(x) = h_\mu(S'),
\]

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

\[
h_\mu(S') = -\lim_{\varepsilon \to 0} E_x^\varepsilon \limsup_{T \to \infty} \frac{1}{T} \ln Q_x^\varepsilon(T, \rho, \delta)
\]

\[
= -\lim_{\varepsilon \to 0} \int_{M} \limsup_{T \to \infty} \frac{1}{T} \ln Q_x^\varepsilon(T, \rho, \delta) m(dx),
\]

where \( E_x^\varepsilon \) is the expectation for the process \( x^\varepsilon_t \) starting at \( x \).
REFERENCES


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

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