Annals of Mathematics, **122** (1985), 509–539

The metric entropy of diffeomorphisms Part I: Characterization of measures satisfying Pesin's entropy formula

By F. LEDRAPPIER and L.-S. YOUNG*

This is the first article in a two-part series containing some results in smooth ergodic theory. We begin by giving an overview of these results. Let M be a compact Riemannian manifold, let $f: M \to M$ be a diffeomorphism, and let m be an f-invariant Borel probability measure on M. There are various ways of measuring the complexity of the dynamical system generated by iterating f. Kolmogorov and Sinai introduced the notion of *metric entropy*, written $h_m(f)$. This is a purely measure-theoretic invariant and has been studied a good deal in abstract ergodic theory (see e.g. [Ro 2]). A more geometric way of measuring chaos is to estimate the exponential rate at which nearby orbits are separated. These rates of divergence are given by the growth rates of Df^n (the derivative of f composed with itself n times). They are called the *Lyapunov exponents* of f and are denoted in this paper by $\{\lambda_i(x): x \in M,\ i = 1, \ldots, r(x)\}$. (See (1.1) for precise definitions.)

The relationship between entropy and exponents has been studied before. A well-known theorem of Margulis and Ruelle [Ru 2] says that entropy is always bounded above by the sum of positive exponents; i.e.,

$$(*) \qquad h_m(f) \leqq \int \sum_i \lambda_i^+(x) \dim E_i(x)\, dm(x)$$

where $\dim E_i(x)$ is the multiplicity of $\lambda_i(x)$ and $a^+ = \max(a, 0)$. Pesin shows that $(*)$ is in fact an equality if f is C^2 and m is equivalent to the Riemannian measure on M. This is sometimes known as *Pesin's formula* [P 2].

Our aim here is to further the study of relations of this type. In Part I we identify those measures for which equality is attained in $(*)$ by their geometric properties. Part II is mainly devoted to proving a formula that is valid for all

*Part of this work was done while the authors were visiting the Mathematical Sciences Research Institute of Berkeley, California. Partially supported by NSF Grant No. MCS 8120790 and AFOSR-83-0265.

invariant measures. This generalized formula contains, in some sense, the above mentioned results of Margulis, Ruelle and Pesin. It involves the notion of dimension and leads to certain volume estimates. These results are announced in [LY].

From here on our discussion will be confined to the subject of Part I.

We attempt to give a brief history leading to this problem. Recall that in the ergodic theory of Anosov diffeomorphisms or of Axiom A attractors, there is an invariant measure that is characterized by each of the following properties:

(1) Equality holds in $(*)$. (In the literature such a measure is sometimes referred to as the equilibrium state of a certain function connected with the derivative of f.)

(2) Its conditional measures on unstable manifolds are absolutely continuous with respect to Lebesgue.

(3) Lebesgue a.e. point in an open set is generic with respect to this measure.

(4) This measure is approximable by measures that are invariant under suitable stochastic perturbations.

Each one of these properties has been shown to be significant in its own right, but perhaps more striking is the fact that they are all equivalent to one another. Many of these ideas are due to Sinai, Bowen and Ruelle. For further information and details we refer the reader to [A], [B], [Ki], [Ru 1], [S 1], [S 2] and [S 3].

At about the same time that progress was being made on uniformly hyperbolic systems, Oseledec [O] proved an ergodic theorem for products of matrices paving the way for analyzing dynamical systems of more general types. Pesin then set up the machinery for translating this linear theory of Lyapunov exponents into non-linear results in neighborhoods of typical trajectories [P 1]. Using these new tools he began to develop an ergodic theory for *arbitrary* diffeomorphisms preserving a measure equivalent to Lebesgue measure [P 2]. (The entropy formula we alluded to earlier is among these first results.) Part of his theory has since been extended and applied to dynamical systems preserving only a *Borel* measure. (See e.g. [Ka] and [Ru 3]; see also [M].)

In view of these developments, it was natural to ask if some of the major results for uniformly hyperbolic systems would remain valid in the more general framework of all C^2 diffeomorphisms. In particular, it was conjectured that properties (1) and (2) above were equivalent. That is, given a diffeomorphism preserving a Borel probability measure m, is it the case that Pesin's formula f holds if and only if m has absolutely continuous conditional measures on unstable manifolds? It is to this conjecture that we shall address ourselves in Part I.

Results partially confirming this conjecture were obtained earlier. That (2) implies (1) is an extension of Pesin's theorem and was proved in [LS]. The reverse implication was proved by the first author [L] under the additional stipulation that the system be at least nonuniformly hyperbolic. We now remove this assumption, confirming the above conjecture in full generality.

To carry out this last step, we have found it necessary to consider explicitly the role played by zero exponents. Indeed, a good portion of our proof consists of an attempt to obtain some control over these nonhyperbolic parts of the dynamical system.

This paper proceeds as follows: Definitions and precise statements of results are given in Section 1. In Section 2 we discuss the estimates and constructions associated with partial nonuniform hyperbolicity. Two partitions are described in Section 3. They are used to estimate the various entropies. Section 4 contains some technical lemmas. These together with all the previous constructions are used in Section 5 to prove the main proposition. The proofs of the theorems are then completed in Section 6.

It is our pleasure to thank our families and a long list of friends and colleagues whose interest and support contributed to the existence of this manuscript. The first author wishes also to acknowledge the hospitality of the University of Maryland.

Standing hypotheses for the entire paper

A. M is a C^∞ compact Riemannian manifold without boundary;
B. f is a C^2 diffeomorphism of M onto itself;
C. m is an f-invariant Borel probability measure on M.

1. Definitions and statements of results

(1.1) For $x \in M$, let T_xM denote the tangent space to M at x. The point x is said to be *regular* if there exist numbers $\lambda_1(x) > \cdots > \lambda_{r(x)}(x)$ and a decomposition of the tangent space at x into $T_xM = E_1(x) \oplus \cdots \oplus E_{r(x)}(x)$ such that for every tangent vector $v \neq 0 \in E_i(x)$,

$$\lim_{n \to \pm\infty} \frac{1}{n} \log \|Df_x^n v\| = \lambda_i(x) \quad \text{and}$$

$$\lim_{n \to \pm\infty} \frac{1}{n} \log |\operatorname{Jac}(Df_x^n)| = \sum_{i=1}^{r(x)} \lambda_i(x) \dim E_i(x).$$

By a theorem of Oseledec [O], the set Γ' of regular points is a set of full measure. The numbers $\lambda_i(x)$, $i = 1, \dots, r(x)$, are called the *Lyapunov exponents* of f at x; $\dim E_i(x)$ is called the multiplicity of $\lambda_i(x)$. The functions $x \mapsto r(x)$, $\lambda_i(x)$ and $\dim E_i(x)$ are invariant along orbits, and so are constant almost everywhere if m is ergodic.

(1.2) Define

$$E^s(x) = \bigoplus_{\lambda_i < 0} E_i(x),$$

$$E^c(x) = E_{i_0}(x) \quad \text{where } \lambda_{i_0}(x) = 0,$$

$$E^u(x) = \bigoplus_{\lambda_i > 0} E_i(x) \quad \text{and}$$

$$W^s(x) = \left\{ y \in M \colon \limsup_{n \to \infty} \frac{1}{n} \log d(f^n x, f^n y) < 0 \right\},$$

where d is the Riemannian metric on M. The set $W^s(x)$ is called the *stable manifold* at x. For $x \in \Gamma'$, if $\dim E^s(x) \neq 0$, then $W^s(x)$ is an immersed submanifold of M of class C^2, tangent at x to $E^s(x)$. The collection $\{W^s(x),\ x \in \Gamma'\}$ is sometimes referred to as the "stable foliation" of f. The *unstable manifold* at x, denoted by $W^u(x)$, and the "unstable foliation" are defined analogously using f^{-1} instead of f. (See [Ru 3] or [FHY] for more details; see also § 2.2 and 4.2.) If W is an immersed submanifold of M, then it inherits a Riemannian structure from M. We denote the corresponding Riemannian measure on W by μ_W.

(1.3) Let $\mathscr{B}$ be the Borel σ-algebra on M completed with respect to m. Then $(M, \mathscr{B}, m)$ is a Lebesgue space; i.e., it is isomorphic to [0, 1] with Lebesgue measure union a countable number of atoms. A *measurable partition* ξ of M is a partition of M such that, up to a set of measure zero, the quotient space M/ξ is separated by a countable number of measurable sets (see [Ro 1]). The quotient space of a Lebesgue space with its inherited probability space structure is again a Lebesgue space. An important property of measurable partitions is that associated with each ξ, there is a canonical system of conditional measures: That is, for every x in a set of full m-measure, there is a probability measure m_x^ξ defined on $\xi(x)$, the element of ξ containing x. These measures are uniquely characterized (up to sets of m-measure 0) by the following properties: If $\mathscr{B}_\xi$ is the sub-σ-algebra of $\mathscr{B}$ whose elements are unions of elements of ξ, and $A \subset M$ is a measurable set, then $x \mapsto m_x^\xi(A)$ is $\mathscr{B}_\xi$-measurable and $m(A) = \int m_x^\xi(A) m(dx)$.

(1.4) Let ξ be a measurable partition of M.

Definition 1.4.1. We say that ξ is *subordinate* to the W^u-foliation if for m-a.e. x, we have

1. $\xi(x) \subset W^u(x)$ and
2. $\xi(x)$ contains a neighborhood of x open in the submanifold topology of $W^u(x)$.

Note that in general the partition into distinct W^u-manifolds is not a measurable partition and that in order for the notion of conditional measures on unstable manifolds to make sense it is necessary to work with measurable partitions subordinate to W^u.

Definition 1.4.2. We say that m has *absolutely continuous conditional measures on unstable manifolds* if for every measurable partition ξ subordinate to W^u, m_x^ξ is absolutely continuous with respect to $\mu_{W^u(x)}$ for a.e. x.

(1.5) THEOREM A. *Let f: $M \hookleftarrow$ be a C^2 diffeomorphism of a compact Riemannian manifold M preserving a Borel probability measure m. Then m has absolutely continuous conditional measures on unstable manifolds if and only if*

$$h_m(f) = \int \sum_i \lambda_i^+(x) \dim E_i(x) m(dx)$$

where $a^+ = \max(a, 0)$.

We prove the "if" part of Theorem A in Part I. The reverse implication is essentially due to Sinai and is proved in precise form in [LS]. (See [M] for an alternate approach. This result also follows from Part II.)

Remark. We show in fact that when the entropy formula in Theorem A is satisfied, the densities $dm_x^\xi/d\mu_{W^u(x)}$ are given by strictly positive functions that are C^1 along unstable manifolds. (See Corollary 6.2.)

(1.6) Define $\mathscr{B}^u$ to be the sub-σ-algebra of $\mathscr{B}$ whose elements are unions of entire W^u-manifolds; $\mathscr{B}^s$ is defined analogously. Recall that the *Pinsker σ-algebra* of f: $(M, \mathscr{B}, m) \hookleftarrow$ is the sub-σ-algebra of $\mathscr{B}$ consisting of sets A such that if $\alpha = \{A, M - A\}$, then $h_m(f, \alpha) = 0$.

THEOREM B. *Let f: $M \hookleftarrow$ be a C^2 diffeomorphism of a compact Riemannian manifold preserving a Borel probability measure m. Then*

$$\mathscr{B}^u \stackrel{\circ}{=} \mathscr{B}^s \stackrel{\circ}{=} \text{the Pinsker } \sigma\text{-algebra of } f.$$

For sub-σ-algebras $\mathscr{B}_1, \mathscr{B}_2 \subset \mathscr{B}$, $\mathscr{B}_1 \stackrel{\circ}{=} \mathscr{B}_2$ means that for every $A_1 \in \mathscr{B}_1$ one has $A_2 \in \mathscr{B}_2$ such that $m(A_1 \Delta A_2) = 0$ and vice versa.

Theorem B was shown to be true for smooth invariant measures by Pesin [P].

We have chosen to prove Theorems A and B by reducing the problems to their respective ergodic cases (see §6). While not at all essential, this line of approach simplifies the presentation, especially where notation is concerned. Thus along with the standing hypotheses stated at the beginning of this paper, we now declare the following

Additional hypothesis for Sections 2–5:

D. m is ergodic.

2. Lyapunov charts and related constructions

We let

$$\begin{aligned} \lambda^+ &= \min\{\lambda_i, \lambda_i > 0\}, \\ \lambda^- &= \max\{\lambda_i, \lambda_i < 0\}, \\ u &= \dim E^u, \\ c &= \dim E^c, \\ s &= \dim E^s \end{aligned}$$

and assume that $u > 0$.

(2.1) *Lyapunov charts.* In this subsection we summarize some results from Pesin theory. Our formulation differs slightly from that in [P1]. See the appendix of this paper for more details.

As always, we let d be the Riemannian metric on M. For

$$(x, y, z) \in \mathbf{R}^u \times \mathbf{R}^c \times \mathbf{R}^s,$$

we define

$$|(x, y, z)| = \max\{|x|_u, |y|_c, |z|_s\}$$

where $|\cdot|_u$, $|\cdot|_c$ and $|\cdot|_s$ are the Euclidean norms on $\mathbf{R}^u$, $\mathbf{R}^c$ and $\mathbf{R}^s$ respectively. The closed disk in $\mathbf{R}^u$ of radius ρ centered at 0 is denoted by $R^u(\rho)$ and $R(\rho) = R^u(\rho) \times R^c(\rho) \times R^s(\rho)$.

Let $0 < \varepsilon < \lambda^+/100, \ -\lambda^-/100$ be given. We shall define in a nonautonomous way a change of coordinates in some neighborhood of each regular point. The size of the neighborhood, the local chart and the estimates will vary with $x \in \Gamma'$. First there is a measurable function l: $\Gamma' \to [1, \infty)$ such that $l(f^{\pm}x) \leq e^{\varepsilon}l(x)$. Then there is an embedding Φ_x: $R(l(x)^{-1}) \to M$ with the following properties:

i) $\Phi_x 0 = x$; $D\Phi_x(0)$ takes $\mathbf{R}^u$, $\mathbf{R}^c$ and $\mathbf{R}^s$ to $E^u(x)$, $E^c(x)$ and $E^s(x)$ respectively:

ii) Let $\tilde{f}_x = \Phi_{fx}^{-1} \circ f \circ \Phi_x$ be the connecting map between the chart at x and the chart at fx, defined wherever it makes sense, and let $\tilde{f}_x^{-1} = \Phi_{f^{-1}x}^{-1} \circ f^{-1} \circ \Phi_x$ be defined similarly. Then

$$e^{\lambda^+ - \varepsilon}|v| \le |D\tilde{f}_x(0)v| \qquad \text{for } v \in \mathbf{R}^u,$$

$$e^{-\varepsilon}|v| \le |D\tilde{f}_x(0)v| \le e^{\varepsilon}|v| \qquad \text{for } v \in \mathbf{R}^c$$

and

$$|D\tilde{f}_x(0)v| \le e^{\lambda^- + \varepsilon}|v| \qquad \text{for } v \in \mathbf{R}^s.$$

iii) If $L(g)$ denotes the Lipschitz constant of the function g, then

$$L\big(\tilde{f}_x - D\tilde{f}_x(0)\big) \le \varepsilon,$$

$$L\Big(\tilde{f}_x^{-1} - D\tilde{f}_x^{-1}(0)\Big) \le \varepsilon$$

and

$$L\big(D\tilde{f}_x\big), L\Big(D\tilde{f}_x^{-1}\Big) \le l(x).$$

iv) For all $z, z' \in R(l(x)^{-1})$, we have

$$K^{-1}d\big(\Phi_x z, \Phi_x z'\big) \le |z - z'| \le l(x)d\big(\Phi_x z, \Phi_x z'\big)$$

for some universal constant K.

It follows from ii) and iii) that there is a number $\lambda > 0$ depending on ε and the exponents such that for all $x \in \Gamma'$, $|\tilde{f}_x z| \le e^{\lambda}|z|$ for all $z \in R(e^{-\lambda-\varepsilon}l(x)^{-1})$. In particular, $\tilde{f}_x R(e^{-\lambda-\varepsilon}l(x)^{-1}) \subset R(l(fx)^{-1})$.

From here on, we shall refer to any system of local charts $\{\Phi_x,\ x \in \Gamma'\}$ satisfying i)–iv) as (ε, l)-charts and λ will be as above.

(2.2) *Local unstable manifolds and center unstable sets.* For very small $\varepsilon > 0$, let $\{\Phi_x,\ x \in \Gamma'\}$ be a system of (ε, l)-charts. Sometimes it is necessary to reduce the size of our charts. Let $0 < \delta \le 1$ be a reduction factor. For $x \in \Gamma'$, define

$$S_\delta^{cu}(x) = \Big\{z \in R\big(l(x)^{-1}\big) \colon \Big|\Phi_{f^{-n}x}^{-1} \circ f^{-n} \circ \Phi_x z\Big| \le \delta l(f^{-n}x)^{-1} \forall n \ge 0\Big\};$$

that is, $\Phi_x S_\delta^{cu}(x)$ consists of those points in M whose backward orbit stays (well) inside the domains of the charts at $f^{-n}x$ for all $n \ge 0$. It is called the *center unstable set* of f at x associated with the charts $\{\Phi_x\}$ and reduction factor δ. On

$S_\delta^{cu}(x)$, we have

$$\tilde{f}_x^{-n} \stackrel{def}{=} \Phi_{f^{-n}x}^{-1} \circ f^{-n} \circ \Phi_x = \tilde{f}_{f^{-n+1}x}^{-1} \circ \cdots \circ \tilde{f}_{f^{-1}x}^{-1} \circ \tilde{f}_x^{-1}.$$

We next introduce the *local unstable manifold* at x associated with $\{\Phi_x\}$ and δ. This is defined to be the component of $W^u(x) \cap \Phi_x R(\delta l(x)^{-1})$ that contains x. The Φ_x^{-1}-image of this set in the x-chart is denoted by $W_{x,\delta}^u(x)$.

PROPOSITION 2.2.1. *Let $\{\Phi_x, x \in \Gamma'\}$ be a system of (ε, l)-charts.*

A. *If $0 < \delta \leq 1$ and $x \in \Gamma'$, then*

i) *$W_{x,\delta}^u(x)$ is the graph of a function*

$$g_x \colon R^u\left(\delta l(x)^{-1}\right) \to R^{c+s}\left(\delta l(x)^{-1}\right)$$

with $g_x(0) = 0$ and $\|Dg_x\| \leq \frac{1}{3}$;

ii) $W_{x,\delta}^u(x) \subset S_\delta^{cu}(x)$.

B. *If $0 < \delta \leq e^{-\lambda-\varepsilon}$ (where λ is as in (2.1)) and $x \in \Gamma'$, then*

$$\tilde{f}_x W_{x,\delta}^u(x) \cap R\left(\delta l(fx)^{-1}\right) = W_{fx,\delta}^u(fx).$$

The proofs of these assertions are standard in unstable manifold theory. We refer the reader to [FHY] for details.

LEMMA 2.2.2. *If $\delta \leq e^{-\lambda-\varepsilon}$, then for almost every $x \in \Gamma'$,*

$$S_\delta^{cu}(x) \cap \Phi_x^{-1} W^u(x) = W_{x,\delta}^u(x).$$

Proof. In view of ii) in Proposition 2.2.1, it suffices to show $S_\delta^{cu}(x) \cap \Phi_x^{-1} W^u(x) \subset W_{x,\delta}^u(x)$. Let $z \in S_\delta^{cu}(x) \cap \Phi_x^{-1} W^u(x)$ and let d^u denote Riemannian distance along W^u-manifolds. Since $\Phi_x z \in W^u(x)$, $d^u(f^{-n}\Phi_x z, f^{-n}x) \to 0$ as $n \to \infty$. But for recurrence reasons, $l(f^{-n}x) \nrightarrow 0$ as $n \to \infty$ for almost every $x \in \Gamma'$. This implies that for almost every $x \in \Gamma'$ there is some $k \geq 0$ such that $\tilde{f}_x^{-k} z \in W_{f^{-k}x,\delta}^u(f^{-k}x)$. Let $k = k(x)$ be the smallest nonnegative integer for which this happens. If $k > 0$, then by Proposition 2.2.1 B, $\tilde{f}_x^{-k+1} z \notin R(\delta l(f^{-k+1}x)^{-1})$, which contradicts $z \in S_\delta^{cu}(x)$. So $k = 0$, or equivalently, $z \in W_{x,\delta}^u(x)$. □

Consider now $y \in \Gamma' \cap \Phi_x S_\delta^{cu}(x)$ where $\delta \leq \frac{1}{4}$. Let $W_{x,2\delta}^u(y)$ be the ϕ_x^{-1}-image of the component of $W^u(y) \cap \Phi_x R(2\delta l(x)^{-1})$. Then $\Phi_x W_{x,2\delta}^u(y)$ contains an open neighborhood of y in $W^u(y)$ and is also referred to as a local unstable manifold at y (although in general $\Phi_y W_{y,\delta}^u(y) \neq \Phi_x W_{x,\delta}^u(y)$). A reduction factor of $\leq \frac{1}{4}$ is taken because working in $f^{-n}x$-charts we cannot control the unstable manifolds of points whose backward orbits come too close to the boundary of $\Phi_{f^{-n}x} R(l(f^{-n}x)^{-1})$. Another technical nuisance is that $\left|\Phi_{f^{-n}x}^{-1} f^{-n} y\right| \neq 0$. Aside from these, we have the analog of (2.2.1) and (2.2.2) for $W^u_{x,2\delta}(y)$ which we state below. The proofs are almost identical to the corresponding ones above.

LEMMA 2.2.3. *Let $\{\Phi_x\}$ be (ε, l)-charts as usual.*

A. *Let $0 < \delta \leq \frac{1}{4}$. Then for every $x \in \Gamma'$ and $y \in \Gamma' \cap S^{cu}_\delta(x)$,*

i) *$W^u_{x,2\delta}(y)$ is the graph of a function*

$$g_{x,y}\colon R^u\big(2\delta l(x)^{-1}\big) \to R^{c+s}\big(2\delta l(x)^{-1}\big)$$

with $\|Dg_{x,y}\| \leq \frac{1}{3}$;

ii) *$W^u_{x,2\delta}(y) \subset S^{cu}_{4\delta}(x)$.*

B. *Let $\delta \leq \min(\frac{1}{4}, \frac{1}{2}e^{-\lambda-\varepsilon})$. For almost every $x \in \Gamma'$, if $y \in \Gamma' \cap S^{cu}_\delta(x)$ and $fy \in S^{cu}_\delta(fx)$, then*

i) *$\tilde{f}_x W^u_{x,2\delta}(y) \cap R(2\delta l(fx)^{-1}) \subset W^u_{fx,2\delta}(fy)$,*

ii) *$S^{cu}_{2\delta}(x) \cap \Phi_x^{-1}W^u(y) \subset W^u_{x,2\delta}(y) \subset S^{cu}_{4\delta}(x) \cap \Phi_x^{-1}W^u(y)$.*

We remark that in general $S^{cu}_\delta(x)$ is a rather messy set. Among other things we think of it as containing pieces of local unstable manifolds (see Lemma 2.2.3, A.ii)). In the case where none of the exponents are zero, $S^{cu}_\delta(x)$ is equal to $W^u_{x,\delta}(x)$.

(2.3) *More estimates.* We list here some estimates that will be used in later sections. Let $\varepsilon, l, \{\Phi_x, x \in \Gamma'\}$ and λ be as before. When working in charts, we use z_u to denote the u-coordinate of the point $z \in R(l(x)^{-1})$. Other notations such as z_s and z_{cu} are understood to have obvious meanings as well.

LEMMA 2.3.1. *Let $\delta \leq e^{-\lambda-\varepsilon}$ and let $x \in \Gamma'$. Then*

(a) *If $z, z' \in R(\delta l(x)^{-1})$ and $|z - z'| = |z_u - z'_u|$, then*

$$\begin{aligned}|\tilde{f}_x z - \tilde{f}_x z'| &= |(\tilde{f}_x z)_u - (\tilde{f}_x z')_u| \\ &\geq e^{\lambda^+ - 2\varepsilon}|z - z'|;\end{aligned}$$

(b) *If u in (a) is replaced by cu, then the conclusion holds with λ^+ replaced by 0;*

(c) *If $z, z' \in S^{cu}_\delta(x)$, then*

$$|\tilde{f}_x^{-1} z - \tilde{f}_x^{-1} z'| \leq e^{2\varepsilon}|z - z'|.$$

Proof. The proofs of (a) and (b) are direct applications of properties ii) and iii) in (2.1). We prove (c): First we claim that $|z_{cu} - z'_{cu}| = |z - z'|$. Suppose not. Then applying (a) to $\tilde{f}_x^{-1}$, we have

$$\begin{aligned}\left|\tilde{f}_x^{-1} z - \tilde{f}_x^{-1} z'\right| &= \left|\left(\tilde{f}_x^{-1} z\right)_s - \left(\tilde{f}_x^{-1} z'\right)_s\right| \\ &\geq e^{-\lambda^- - 2\varepsilon}|z - z'|.\end{aligned}$$

Inductively, this gives

$$\begin{aligned}|\tilde{f}_x^{-n}z - \tilde{f}_x^{-n}z'| &= |(\tilde{f}_x^{-n}z)_s - (\tilde{f}_x^{-n}z')_s| \\ &\geq e^{-(\lambda^- + 2\varepsilon)n}|z - z'|\end{aligned}$$

for all $n \geq 0$, which forces one of the points $\tilde{f}_x^{-n}z$ or $\tilde{f}_x^{-n}z'$ to leave the chart, contradicting $z, z' \in S_\delta^{cu}(x)$. Now this argument also applies to $\tilde{f}_x^{-1}z$ and $\tilde{f}_x^{-1}z'$, since they belong in $S_\delta^{cu}(f^{-1}x)$. It then follows from (b) that

$$|z - z'| = |z_{cu} - z'_{cu}| \geq e^{-2\varepsilon}|\tilde{f}_x^{-1}z - \tilde{f}_x^{-1}z'|,$$

which is the desired conclusion. □

The next lemma involves some estimates in the charts at x and fx where both of these points belong in Γ'.

LEMMA 2.3.2. *Assume* $\delta \leq \min(\frac{1}{4}, \frac{1}{2}e^{-\lambda-\varepsilon})$. *Let* $y \in S_\delta^{cu}(x)$, $z = (\{0\} \times \mathbf{R}^{c+s}) \cap W_{x,2\delta}^u(y)$ *and* $z' = (\{0\} \times \mathbf{R}^{c+s}) \cap \tilde{f}_x W_{x,2\delta}^u(y)$. *Then*

$$|z'| \leq e^{3\varepsilon}|z|.$$

Proof. Since $W_{x,2\delta}^u(y)$ is the graph of a function $g_{x,y}$ with $\|Dg_{x,y}\| \leq \frac{1}{3}$, the slope of $\tilde{f}_x W_{x,2\delta}^u(y)$ is ≤ 1. This gives

$$|z'| \leq |(\tilde{f}_x z)_{cs}| + |(\tilde{f}_x z)_u|.$$

From properties ii) and iii) of (2.1), it follows that $|(\tilde{f}_x z)_{cs}| \leq (e^\varepsilon + \varepsilon)|z|$ and $|(\tilde{f}_x z)_u| \leq \varepsilon|z|$, so that $|z'| \leq e^{3\varepsilon}|z|$. □

(2.4) *Partitions adapted to Lyapunov charts.* In order to make use of the geometry of Lyapunov charts in the calculation of entropy, it is convenient to have partitions whose elements lie in charts. If $\mathscr{P}$ is a partition of M, write $\mathscr{P}^+ = \bigvee_{n=0}^\infty f^n\mathscr{P}$. Let ε, l, $\{\Phi_x, x \in \Gamma'\}$ and λ be as before, and let $0 < \delta \leq 1$ be a reduction factor.

Definition 2.4.1. A measurable partition $\mathscr{P}$ is said to be *adapted* to $(\{\Phi_x\}, \delta)$ if for almost every $x \in \Gamma'$, $\mathscr{P}^+(x) \subset \Phi_x S_\delta^{cu}(x)$.

LEMMA 2.4.2. *Given* $\{\Phi_x\}$ *and* $0 < \delta \leq 1$, *there is a finite entropy partition* $\mathscr{P}$ *such that* $\mathscr{P}$ *is adapted to* $(\{\Phi_x\}, \delta)$.

Proof. We outline the construction of $\mathscr{P}$ using an idea of Mañé's [M]. Fix some $l_0 > 0$ and let $\Lambda \subset \Gamma' \cap \{l(x) \leq l_0\}$. Assume that $m\Lambda > 0$. For $x \in \Lambda$, let $r(x)$ be the smallest positive integer k such that $f_x^k \in \Lambda$. We define

$\psi \colon \Lambda \to \mathbf{R}^+$ by

$$\psi(x) = \begin{cases} \delta & \text{if } x \notin \Lambda \\ \delta l_0^{-2} e^{-(\lambda+\varepsilon) r(x)} & \text{if } x \in \Lambda. \end{cases}$$

Then ψ is defined almost everywhere on Λ and $\log \psi$ is integrable since $\int_\Lambda r\, dm = 1$. Let $B(x, \rho) = \{ y \in M \colon d(x, y) < \rho \}$. By Lemma 2 in [M], there is a partition $\mathscr{P}$ with $H_m(\mathscr{P}) < \infty$ such that $\mathscr{P}(x) \subset B(x, \psi(x))$ for almost every x. We claim that this $\mathscr{P}$ is adapted to $(\{\Phi_x\}, \delta)$. In fact, we will show that $\mathscr{P}^+(x) \subset \Phi_x R(\delta l(x)^{-1})$ for almost every $x \in \bigcup_{n \geq 0} f^n \Lambda$.

First consider $x \in \Lambda$. By choice of $\mathscr{P}$, we have $\mathscr{P}^+(x) \subset \mathscr{P}(x) \subset B(x, \psi(x))$ which is contained in $\Phi_x R(\delta l(x)^{-1})$ because $\psi(x) l(x) = \delta l_0^{-2} e^{-(\lambda+\varepsilon) r(x)} l(x) \leq \delta l(x)^{-1}$. Suppose now that $x \notin \Lambda$ and $n > 0$ is the smallest positive integer such that $f^{-n} x \in \Lambda$. Then $f^{-n} \mathscr{P}^+(x) \subset \mathscr{P}^+(f^{-n}x) \subset B(f^{-n}x, \psi(f^{-n}x))$. Now

$$\begin{aligned} f^n B(f^{-n}x, \psi(f^{-n}x)) &\subset \Phi_x \tilde{f}^n_{f^{-n}x} R\Big(\delta l(f^{-n}x)^{-1} e^{-(\lambda+\varepsilon) r(f^{-n}x)}\Big) \\ &\subset \Phi_x R\Big(\delta l(f^{-n}x)^{-1} e^{-(\lambda+\varepsilon) r(f^{-n}x)} e^{\lambda n}\Big) \\ &\subset \Phi_x R\Big(\delta l(x)^{-1}\Big) \end{aligned}$$

since $n \leq r(f^{-n}x)$. Note that this computation makes sense because for every $1 \leq k < n$, $\tilde{f}^k_{f^{-n}x} R(\delta l(f^{-n}x)^{-1} e^{-(\lambda+\varepsilon) r(f^{-n}x)}) \subset R(l(f^{-n+k}x)^{-1} e^{-(\lambda+\varepsilon)})$. This completes the proof. □

3. Construction of two partitions

(3.1) *Increasing partitions subordinate to the W^u-foliation.* Given two partitions ξ_1 and ξ_2 of M, we say that ξ_1 refines ξ_2 ($\xi_1 > \xi_2$) if at m-a.e. $x \in M$, $\xi_1(x) \subset \xi_2(x)$. A partition is said to be *increasing* if $\xi > f\xi$.

In this subsection we describe a family of increasing partitions that are subordinate to the unstable foliation. Partitions of this type were used by Sinai [S1] to study uniformly hyperbolic systems and are discussed in detail in [LS].

LEMMA 3.1.1. *There exist measurable partitions with the following properties*:

1) *ξ is an increasing partition subordinate to W^u*;
2) *$\bigvee_{n=0}^{\infty} f^{-n}\xi$ is the partition into points*;
3) *the biggest σ-algebra contained in $\bigcap_{n=0}^{\infty} f^n \xi$ is $\mathscr{B}^u$.*

The construction we sketch below not only proves Lemma 3.1.1 but produces partitions with certain additional properties that will be useful later.

Outline of construction. Let $\{\Phi_x, x \in \Gamma'\}$ be a system of (ε, l)-local charts and let l_0 be a number such that $m(l \leq l_0) > 0$. We claim that there is a measurable set S with the following properties:

(a) $mS > 0$;

(b) S is the disjoint union of a continuous family of embedded disks $\{D_\alpha\}$, where each D_α is an open subset of $\Phi_{x_\alpha} W^u_{x_\alpha, 1}(x_\alpha)$ for some $x_\alpha \in \{l \leq l_0\}$;

(c) For almost every $x \in M$, there is an open neighborhood U_x of x in $W^u(x)$ such that for each $n \geq 0$, either $f^{-n}U_x \cap S = \varnothing$ or $f^{-n}U_x \subset D_\alpha$ for some α;

(d) (This requirement is irrelevant for proving Lemma 3.1.1.) There is a number γ such that:

i) The d^u-diameter of every D_α in S is less than γ and

ii) If $x, y \in S$ are such that $y \in W^u(x)$ and $d^u(x, y) > \gamma$, then x and y lie on distinct D_α-disks.

The existence of an S satisfying (a), (b) and (c) is proved in [LS] and will not be repeated here. Property (d) is easy to arrange by cutting down the d^u-size of the disks in [LS]. Let $\hat{\xi}$ be the partition of M defined by

$$\hat{\xi}(x) = \begin{cases} D_\alpha & \text{if } x \in D_\alpha \\ M - S & \text{if } x \notin S; \end{cases}$$

then $\xi = \hat{\xi}^+$ is the partition we desire. It is easy to verify that it has the properties stated in Lemma 3.1.1. □

The partitions whose construction we just outlined have the following alternate characterization: There is a set S satisfying (a)–(d) such that if $\sigma = \bigvee_{n=0}^{\infty} f^n\{S, M - S\}$, then for every $x \in M$, $y \in \xi(x)$ if and only if $y \in \sigma(x)$ and $d^u(f^{-n}x, f^{-n}y) \leq \gamma$ whenever $f^{-n}x \in S$.

For measurable partitions η_1 and η_2, let $H_m(\eta_1|\eta_2)$ denote the mean conditional entropy of η_1 given η_2. Note that if η is an increasing partition, then $h_m(f, \eta) = H_m(\eta|f\eta)$.

LEMMA 3.1.2. *Let ξ_1 and ξ_2 be partitions constructed in the proof of Lemma 3.1.1. Then*

$$h_m(f, \xi_1) = h_m(f, \xi_2).$$

Proof. It suffices to show $h(f, \xi_1 \vee \xi_2) = h(f, \xi_1)$. For every $n \geq 1$, we have

$$\begin{aligned} h(f, \xi_1 \vee \xi_2) &= h(f, \xi_1 \vee f^n\xi_2) \\ &= H\big(\xi_1 \vee f^n\xi_2 | f\xi_1 \vee f^{n+1}\xi_2\big) \\ &= H\big(\xi_1 | f\xi_1 \vee f^{n+1}\xi_2\big) + H\big(\xi_2 | f\xi_2 \vee f^{-n}\xi_1\big). \end{aligned}$$

As $n \to \infty$, the second term goes to 0 since $f^{-n}\xi_1$ generates. We claim that $H(\xi_1|f\xi_1 \vee f^{n+1}\xi_2) \to H(\xi_1|f\xi_1)$. Clearly, $H(\xi_1|f\xi_1 \vee f^{n+1}\xi_2) \le H(\xi_1|f\xi_1)$ for all $n \ge 0$. Let $D_n = \{x\colon (f\xi_1)(x) \subset (f^n\xi_2)(x)\}$. Since for almost every x, the d^u-diameter of $\xi_1(x)$ is finite and d^u-diam$(f^{-n}\xi_1)(x) \downarrow 0$ as $n \to \infty$, we have $mD_n \to 1$. Thus for large enough n, there is a set D_n with measure arbitrarily close to 1 such that restricted to D_n, $f\xi_1 \vee f^n\xi_2 = f\xi_1$. This proves

$$\lim H\big(\xi_1|f\xi_1 \vee f^{n+1}\xi_2\big) \ge H(\xi_1|f\xi_1). \qquad \square$$

(3.2) *Two useful partitions.* Let $\{\Phi_x, x \in \Gamma'\}$ be a system of (ε, l)-charts, let λ be as before and let ξ be an increasing partition subordinate to W^u constructed as in the proof of Lemma 3.1.1, with l_0, S and γ having the same meaning as in (3.1). Let $\delta \le \min(\frac{1}{4}, \frac{1}{2}e^{-\lambda-\varepsilon}, \gamma/2K)$ and let $\mathscr{P}$ be a finite entropy partition adapted to $(\{\Phi_x\}, \delta)$. We require that $\mathscr{P}$ refine $\{S, M - S\}$ and another finite entropy partition to be specified later. Define

$$\eta_1 = \xi \vee \mathscr{P}^+ \quad \text{and}$$
$$\eta_2 = \mathscr{P}^+.$$

These two partitions play central roles in Section 5. We compare their properties:

1) Both η_1 and η_2 are increasing measurable partitions,

2) $\eta_1 > \eta_2$,

3) $\eta_2(x) \subset \Phi_x S^{cu}_\delta(x)$ and $\eta_1(x) \subset \Phi_x W^u_{x,\delta}(x)$ for m-a.e. x,

and

4) $h_m(f, \eta_2) = h_m(f, \mathscr{P})$ and $h_m(f, \eta_1) = H_m(\xi|f\xi)$.

Properties 1) and 2) and the first half of 3) follow from the definitions of η_1 and η_2. The second half of 3) is a consequence of Lemma 2.2.2. The first half of 4) is straightforward. We prove the remaining assertion:

Lemma 3.2.1. $h(f, \eta_1) = H_m(\xi|f\xi)$.

Proof. As in the argument in Lemma 3.1.2, we have

$$\begin{aligned} h(f, \eta_1) &= h(f, \xi \vee f^n\mathscr{P}^+) \\ &= H\big(\xi|f\xi \vee f^{n+1}\mathscr{P}^+\big) + H\big(\mathscr{P}^+|f^{-n}\xi \vee f\mathscr{P}^+\big) \end{aligned}$$

where the first term is $\le H(\xi|f\xi)$ and the second term goes to 0 as $n \to \infty$. Also, using the fact that $H(\mathscr{P}) < \infty$, we have

$$h(f, \eta_1) = h(f, \xi \vee \mathscr{P}) \ge h(f, \xi). \qquad \square$$

(3.3) *Quotient Structure.* Since $\eta_1 > \eta_2$ we can view η_1 restricted to each $\eta_2(x)$ (written $\eta_1|\eta_2(x)$) as a subpartition of $\eta_2(x)$. This subpartition has a simple geometric description in the x-chart: recall that since $\eta_2(x) \subset \Phi_x S^{cu}_\delta(x)$ where

$\delta \leq \frac{1}{4}$, for every $y \in S_\delta^{cu}(x)$, $W_{x,2\delta}^u(y)$ is the graph of a function from $R^u(2\delta l(x)^{-1})$ to $R^{c+s}(2\delta l(x)^{-1})$. The restriction of these graphs to $\Phi_x^{-1}\eta_2(x)$ gives a natural partition of $\Phi_x^{-1}\eta_2(x)$. The next lemma says that this corresponds to $\eta_1|\eta_2(x)$.

LEMMA 3.3.1. *For almost every x and every $y \in \Gamma' \cap \eta_2(x)$,*

$$\Phi_x W_{x,2\delta}^u(y) \cap \eta_2(x) = \eta_1(y).$$

Proof. First consider $z \in \Phi_x W_{x,2\delta}^u(y) \cap \eta_2(x)$. We will show that $z \in \xi(y)$. Since $\mathscr{P}$ refines $\{S, M - S\}$ and $z \in \mathscr{P}^+(y)$, it suffices to show (using the characterization of ξ in (3.1)) that $d^u(f^{-n}y, f^{-n}z) \leq \gamma$ whenever $f^{-n}y \in S$. This is in fact true for all $n \geq 0$, for $|\Phi_x^{-1}y - \Phi_x^{-1}z| \leq 2\delta l(x)^{-1}$ and by Lemma 2.3.1,

$$\left|\tilde{f}_x^{-n}\Phi_x^{-1}y - \tilde{f}_x^{-n}\Phi_x^{-1}z\right| \leq \left|\Phi_x^{-1}y - \Phi_x^{-1}z\right| \quad \text{for all } n \geq 0.$$

Together these imply that $d^u(f^{-n}y, f^{-n}z) \leq K2\delta l(x)^{-1} \leq \gamma$. The reverse containment follows from Lemma 2.2.3, B.ii). □

This lemma allows us to identify the quotient space $\eta_2(x)/\eta_1$ with a subset of $\mathbf{R}^{c+s}$ via $\eta_1(y) \leftrightarrow W_{x,2\delta}^u(y) \cap \{0\} \times \mathbf{R}^{c+s}$. The next lemma tells us that the map $f|f^{-1}(\eta_2(x))$: $f^{-1}(\eta_2(x)) \to \eta_2(x)$ acts like a skew product with respect to this quotient structure.

LEMMA 3.3.2. *For almost every x and every $y \in \Gamma' \cap \eta_2(x)$,*

$$f^{-1}(\eta_1(y)) = \eta_1(f^{-1}y) \cap f^{-1}(\eta_2(x)).$$

Proof. First $f^{-1}(\eta_1(y)) \subset f^{-1}(\eta_2(x))$ because $\eta_1(y) \subset \eta_2(x)$ and $f^{-1}(\eta_1(y)) \subset \eta_1(f^{-1}y)$ because η_1 is an increasing partition. That

$$f\left(\eta_1(f^{-1}y)\right) \cap \eta_2(x) \subset \eta_1(y)$$

follows immediately from Lemma 3.3.1 and Lemma 2.2.3, B.i). □

(3.4) *Transverse metrics.* To use the fact that all the expansion of f occurs *along* the W^u-foliation, we need to show that the map induced by f on $(f^{-1}\eta_2(x))/\eta_1 \to \eta_2(x)/\eta_1$ does not expand distances. To that end we define a metric on the quotient space $\eta_2(x)/\eta_1$ for m-a.e. x. This will be referred to as a *transverse metric.*

As far as we know, "canonical" systems of transverse metrics do not exist.

First we give a point-dependent definition: Let $x \in \Gamma'$. From (2.2) we know that for every $y \in \eta_2(x)$, $W_{x,2\delta}^u(y)$ intersects $\{0\} \times \mathbf{R}^{c+s}$ at exactly one point. We call this point z. For $y' \in \eta_2(x)$, let z' be the corresponding point in

$\{0\} \times \mathbf{R}^{c+s}$. Then

$$d'_x(y, y') \stackrel{\text{def}}{=} |z - z'|.$$

Note that $d'_x(\cdot, \cdot)$ induces a metric on $\eta_2(x)/\eta_1$, but that in general, $d'_x(\cdot, \cdot) \neq d'_{x'}(\cdot, \cdot)$ for $x' \in \eta_2(x)$, $x' \neq x$.

To rectify this situation, we (arbitrarily) choose a reference plane T and standardize all measurements with respect to T. Let S be the set in the construction of ξ (see (3.1)), the partition from which η_1 and η_2 are eventually derived. Let $E \subset S$ be a measurable set with $mE > 0$. Further assumptions on the diameter of E will be given in (4.2). Let τ be the C^2 embedding of a $(c + s)$-dimensional disk into M. We assume that T, the image of τ, is transverse to every D_α in S and intersects D_α in exactly one point if $D_\alpha \cap E \neq \varnothing$. Finally, we require that the partition $\mathscr{P}$ in the definition of η_1 and η_2 refine $\{E, M - E\}$. (See (3.2).)

With this setup, we can now define a metric on $\eta_2(x)/\eta_1$ for every $x \in \bigcup_{n \geq 0} f^n E$. First define a function π: $\bigcup_{n \geq 0} f^n E \to \mathbf{R}^{c+s}$ as follows: For $x \in E \cap D_\alpha$, let

$$\pi(x) = \tau^{-1}\{T \cap D_\alpha\}$$

and in general, let

$$\pi(x) = \pi\left(f^{-n(x)}x\right)$$

where $n(x)$ is the smallest nonnegative integer such that $f^{-n(x)}x \in E$. Then for $x \in \bigcup_{n \geq 0} f^n E$ and y, $y' \in \eta_2(x)$, define

$$d^T_x(y, y') = |\pi y - \pi y'|$$

where $|\cdot|$ denotes Euclidean distance in $\mathbf{R}^{c+s}$

Note that since $\eta_2 = \mathscr{P}^+$, for every $n \geq 0$ either $f^{-n}(\eta_2(x)) \subset E$ or $f^{-n}(\eta_2(x)) \cap E = \varnothing$. Also, when $f^{-n}x \in E$, $f^{-n}(\eta_1(x)) \subset D_\alpha$ for some α. This guarantees that $d^T_x(\cdot, \cdot)$ induces a genuine metric on each $\eta_2(x)/\eta_1$ and that for $x' \in \eta_2(x)$, $d^T_x = d^T_{x'}$.

We comment here on the arbitrariness of our choice of T. It will become clear after (4.2) that for given E, if T and T' are admissible transversals, then d^T_x is uniformly equivalent to $d^{T'}_x$ for m-a.e. x. Furthermore, it is on the equivalence classes of $\{d^T_x\}$, not the metrics themselves, that our estimates in Section 5 depend.

Finally, what we have done here is to represent M/η_1 as a subset of $M/\eta_2 \times \mathbf{R}^{c+s}$, and to define transverse metrics on $\eta_2(x)/\eta_1$ that correspond to Euclidean distance on $\mathbf{R}^{c+s}$. This Euclidean space geometry plays a role in some of our averaging arguments as we shall see in (4.1).

4. Some technical lemmas

(4.1) *A covering lemma and some consequences*. For $x \in \mathbf{R}^n$, let $B(x, r)$ denote the ball of radius r centered at x. All distances are Euclidean in this subsection.

Besicovitch Covering Lemma (BCL, [G]). *Given a set $E \subset \mathbf{R}^n$ and an arbitrary function r: $E \to (0, \infty)$ with $\sup_{x \in E} r(x) < +\infty$, let $\mathscr{A} = \{B(x, r(x)), x \in E\}$. Then there exists a subcover $\mathscr{A}' \subset \mathscr{A}$ such that no x in $\mathbf{R}^n$ lies in more than $c(n)$ elements of $\mathscr{A}'$, $c(n)$ depending only on n.*

Now let μ be a Borel probability measure on $\mathbf{R}^n$. The next two lemmas are standard when μ is Lebesgue. When working with arbitrary finite Borel measures, we use Besicovitch's covering lemma instead of Vitali's lemma. (This, of course, is not new.) Let $g \in L^1(\mu)$ and define

$$g_\delta(x) = \frac{1}{\mu B(x, \delta)} \int_{B(x, \delta)} g \, d\mu.$$

For g positive, we further define

$$g^* = \sup_{\delta > 0} g_\delta \quad \text{and}$$
$$g_* = \inf_{\delta > 0} g_\delta.$$

First we have the maximal lemmas.

Lemma 4.1.1. (a) *For $\lambda \in \mathbf{R}^+$,*

$$\mu(g^* > \lambda) \leq \frac{c(n)}{\lambda} \int g \, d\mu$$

(b) *Let ν be defined by $d\nu = g \, d\mu$. Then for $\lambda \in \mathbf{R}^+$,*

$$\nu(g_* < \lambda) \leq c(n) \lambda.$$

Proof. We give a proof of (a). Part (b) is proved similarly. Let $A = \{g^* > \lambda\}$. For each $x \in A$, choose $\delta(x)$ such that $g_{\delta(x)} > \lambda$; i.e. $\int_{B(x, \delta(x))} g \, d\mu > \lambda \mu B(x, \delta(x))$. Letting $\mathscr{A} = \{B(x, \delta(x)), x \in A\}$ and choosing $\mathscr{A}'$ as in BCL we have

$$\mu(A) \leq \sum_{B \in \mathscr{A}'} \mu(B)$$
$$\leq \sum_{B \in \mathscr{A}'} \frac{1}{\lambda} \int_B g \, d\mu \leq \frac{c(n)}{\lambda} \int_{\mathbf{R}^n} g \, d\mu. \qquad \square$$

Lemma 4.1.2. *Let $g \in L^1(\mu)$. Then $g_\delta \to g$ almost everywhere.*

Proof. This is obvious if g is continuous. Using Part (a) of Lemma 3.1.1 we can show that the set of functions g for which this is true is norm closed in $L^1(\mu)$. □

The next lemma is usually stated slightly differently in the literature. For geometric reasons we average over balls instead of taking conditional expectations with respect to fixed partitions.

LEMMA 4.1.3. *Let (X, μ) be a Lebesgue space and let $\pi: X \to \mathbf{R}^n$ be a measurable map. Disintegrate μ to get a family of probability measures $\{\mu_t\}_{t \in \mathbf{R}^n}$. Let α be a partition of X with $H_\mu(\alpha) < \infty$. For $t \in \mathbf{R}^n$ and $A \in \alpha$, define*

$$g^A(t) = \mu_t(A).$$

Let g_δ^A and g_*^A be functions on $\mathbf{R}^n$ defined as above. Let g, g_δ and g_*: $X \to \mathbf{R}$ be given by

$$g(x) = \sum_{A \in \alpha} \chi_A(x) g^A(\pi x),$$

$$g_\delta(x) = \sum_{A \in \alpha} \chi_A(x) g_\delta^A(\pi x) \quad \textit{and}$$

$$g_*(x) = \sum_{A \in \alpha} \chi_A(x) g_*^A(\pi x).$$

Then $g_\delta \to g$ almost everywhere on X and

$$\int -\log g_* \, d\mu \le H_\mu(\alpha) + \log c + 1$$

where $c = c(n)$ is as in BCL.

Proof. First by Lemma 4.1.2 we have $g_\delta^A \to g^A \mu \circ \pi^{-1}$ a.e. on $\mathbf{R}^n$ and hence $g_\delta \to g$ a.e. on X. Note also that

$$\int -\log g_* \, d\mu = \int_0^\infty \mu(-\log g_* > s) \, ds$$
$$= \int_0^\infty \sum_{A \in \alpha} \mu\left(A \cap \{g_*^A \circ \pi < e^{-s}\}\right) ds.$$

Now

$$\mu\left(A \cap \{g_*^A \circ \pi < e^{-s}\}\right) \le \mu(A).$$

Also,

$$\mu\left(A \cap \{g_*^A \circ \pi < e^{-s}\}\right) \le \int_{\{g_*^A < e^{-s}\}} g^A d(\mu \circ \pi^{-1})$$
$$\le c(n) e^{-s} \quad \text{by Lemma 4.1.1(b)}.$$

Thus

$$\int -\log g_* \, d\mu \le \sum_{A \in \alpha} \int_0^\infty \min(c(n)e^{-s}, \mu(A))\, ds$$
$$\le H_\mu(\alpha) + \log c(n) + 1$$

by a simple calculation. □

Another consequence of BCL is the following classical result (whose proof we omit):

LEMMA 4.1.4. *Let μ be a finite Borel measure on $\mathbf{R}^n$. Then*

$$\inf_{0<\varepsilon\le 1} \frac{\mu B(x,\varepsilon)}{\varepsilon^n} > 0$$

for μ-a.e. x. In particular,

$$\limsup_{\varepsilon\to 0} \frac{\log \mu B(x,\varepsilon)}{\log \varepsilon} \le n.$$

(4.2) *Lipschitz property of local unstable manifolds within center unstable sets.* This is the only part of our construction where it is essential to assume that f is C^2, or at least C^{1+1}, as opposed to $C^{1+\alpha}$ for some $\alpha > 0$. We will be working exclusively in charts and all notations are as in Section 2.

Let $L(\mathbf{R}^u, \mathbf{R}^{c+s})$ denote the set of linear maps from $\mathbf{R}^u$ to $\mathbf{R}^{c+s}$ with norm $\le \frac{1}{3}$. Fix $x \in \Gamma'$ and let $U \subset R(l(x)^{-1})$ be such that $\tilde{f}_x U \subset R(l(fx)^{-1})$. To simplify notation we write $F = \tilde{f}_x$. For $z \in U$, define Ψ_z: $L(\mathbf{R}^u, \mathbf{R}^{c+s}) \hookleftarrow$ by

$$DF_z \text{graph}(v) = \text{graph}(\Psi_z v)$$

where $v \in L(\mathbf{R}^u, \mathbf{R}^{c+s})$. Given g: $U \to L(\mathbf{R}^u, \mathbf{R}^{c+s})$, if Ψg is the function from $F(U)$ to $L(\mathbf{R}^u, \mathbf{R}^{c+s})$ given by

$$\Psi g(Fz) = \Psi_z(g(z)),$$

then Ψg is called the *graph transform* of g by (F, DF). Similar ideas are discussed in [HP], for instance.

In what follows, $L(\cdot)$ denotes the Lipschitz constant of the map.

LEMMA 4.2.1. *Let x and U be as above, and let g: $U \to L(\mathbf{R}^u, \mathbf{R}^{c+s})$ be Lipschitz. Then Ψg is Lipschitz with*

$$L(\Psi g) \le e^{-\lambda^+ + 6\varepsilon} L(g) + 4e^{2\varepsilon} l(x).$$

Proof. By ii) and iii) of (2.1), it is straightforward to verify that for all $y \in U$ and for all $u, v \in L(\mathbf{R}^u, \mathbf{R}^{c+s})$,

$$|\Psi_y u - \Psi_y v| \le e^{-\lambda^+ + 4\varepsilon} |u - v|.$$

Now

$(*)$

$$|\Psi g(Fy) - \Psi g(Fy')| \leq |\Psi_y(g(y)) - \Psi_y(g(y'))| + |\Psi_y(g(y')) - \Psi_{y'}(g(y'))|.$$

The first term of the right-hand side of $(*)$ is $\leq e^{-\lambda^+ + 4\varepsilon}|g(y) - g(y')|$, so that

$$\begin{aligned}\frac{|\Psi_y(g(y)) - \Psi_y(g(y'))|}{|Fy - Fy'|} &\leq e^{-\lambda^+ + 4\varepsilon}\frac{|g(y) - g(y')|}{|y - y'|} \cdot \frac{|y - y'|}{|Fy - Fy'|} \\ &\leq e^{-\lambda^+ + 4\varepsilon} \cdot L(g) \cdot e^{2\varepsilon} \quad \text{by Lemma 2.3.1(c).}\end{aligned}$$

A simple calculation shows that the second term is $\leq 4|DF_y - DF_{y'}|$. By iii) of (2.1), this is $\leq 4l(x)|y - y'|$. Thus

$$\begin{aligned}\frac{|\Psi_y(g(y')) - \Psi_{y'}(g(y'))|}{|Fy - Fy'|} &\leq 4l(x)\frac{|y - y'|}{|Fy - Fy'|} \\ &\leq 4e^{2\varepsilon}l(x).\end{aligned}$$

□

Recall from (2.2) that for $x \in \Gamma'$ and $\delta \leq \frac{1}{4}$, if $y \in S_\delta^{cu}(x)$, then $W_{x,2\delta}^u(y)$ is well defined. A well-known fact from unstable manifold theory is that if $g_0\colon S_\delta^{cu}(x) \to L(\mathbf{R}^u, \mathbf{R}^{c+s})$ is implicitly defined by

$$D\Psi_x(\text{graph } g_0(z)) = E^u(\Psi_x z),$$

then $g_0 = \lim_{n\to\infty}\Psi^n 0$, where Ψ^n is the graph transform by $(\tilde{f}^n_{f^{-n}x}, D\tilde{f}^n_{f^{-n}x})$ and 0 is the zero function from $\tilde{f}_x^{-n}S_\delta^{cu}(x)$ to $L(\mathbf{R}^u, \mathbf{R}^{c+s})$.

Lemma 4.2.2. *Let $x \in \Gamma'$ and $\delta \leq \frac{1}{4}$. Then g_0 as defined above is Lipschitz with*

$$L(g_0) \leq Dl(x)$$

where D is a constant independent of x.

Proof. Letting $0\colon \tilde{f}_x^{-n}S_\delta^{cu}(x) \to L(\mathbf{R}^u, \mathbf{R}^{c+s})$ be the trivial map and using Lemma 4.2.1, we can show inductively that

$$\begin{aligned}L(\Psi^n 0) &\leq 4e^{2\varepsilon}\sum_{i=0}^{n-1} e^{(-\lambda+6\varepsilon)i}l(f^{-i-1}x) \\ &\leq 4l(x)e^{3\varepsilon}\sum_{i=0}^{n-1} e^{(-\lambda^+ +7\varepsilon)i}.\end{aligned}$$

This geometric series converges as $n \to \infty$ since $\varepsilon < \lambda^+/100$. (See (2.1).) The uniform Lipschitz property of $\Psi^n 0$ for all n passes on to g_0. □

We have shown that on $S_\delta^{cu}(x)$, the tangent bundle to $W_{x,2\delta}^u(y)$, $y \in S_\delta^{cu}(x)$, is Lipschitz with Lipschitz constant $\leq Dl(x)$. It follows from this that $\{W_{x,2\delta}^u(y), y \in S_\delta^{cu}(x)\}$ is a Lipschitz lamination, meaning that the Poincaré map between transversals (wherever it make sense) is also Lipschitz with Lipschitz constant proportional to $l(x)$.

We will record this corollary in a convenient form, but first we give further specification on E and T.

Let $E \subset S \cap \{l \leq l_0\}$ be as in (3.4) and have (arbitrarily small) positive measure. We can take T as before, but for the sake of definiteness, let us fix some $w_0 \in E$ and let $\tau = \Phi_{w_0}|\{0\} \times R^{c+s}(l(w_0)^{-1})$. Then T is contained in the $\exp_{w_0}$-image of a neighborhood of 0 in $E^{c+s}(w_0)$. We assume the diameter of E is small enough that for all $x \in E$, $\Phi_x^{-1}w_0 \in R(\frac{1}{4}l(x)^{-1})$ and $\Phi_x^{-1}T$ is the graph of a function from $R^{c+s}(\frac{1}{2}l(x)^{-1})$ to $R^u(\frac{1}{2}l(x)^{-1})$ with slope $\leq 1/100$. This is possible to arrange since $x \mapsto E^{c+s}(x)$ is continuous on $\{l \leq l_0\}$ and all chart estimates are uniform on $\{l \leq l_0\}$.

Lemma 4.2.3. *Let E and T be as above. Then there is a number $N = N(l_0)$ such that for all $x \in E$,*

$$\frac{1}{N} d_x'(\cdot, \cdot) \leq d_x^T(\cdot, \cdot) \leq N d_x'(\cdot, \cdot).$$

Proof. In the chart at x, we define the Poincaré map

$$\theta\colon (\{0\} \times \mathbf{R}^{c+s}) \cap \{W_{x,2\delta}^u(y), y \in S_\delta^{cu}(x)\} \to \Phi_x^{-1}T$$

by sliding along $W_{x,2\delta}^u(y)$. Lemma 4.2.2 tells us that there is a number D' independent of x such that

$$L(\theta), L(\theta^{-1}) \leq D'l(x).$$

Thus if $y, y' \in \eta_2(x)$, and z and z' are respectively the points of intersection of $W_{x,2\delta}^u(y)$ and $W_{x,2\delta}^u(y')$ with $\Phi_x^{-1}T$, then

$$|z - z'|_{\Phi_x^{-1}T} \leq D'd_x'(y, y')$$

where $|\cdot - \cdot|_{\Phi_x^{-1}T}$ denotes distance in $R(l(x)^{-1})$ along the submanifold $\Phi_x^{-1}T$. Therefore we have

$$\begin{aligned} d_x^T(y, y') &\stackrel{\text{def}}{=} |\pi y - \pi y'| = |\Phi_{w_0}^{-1}\Phi_x z - \Phi_{w_0}^{-1}\Phi_x z'| \\ &\leq l_0 K D' d_x'(y, y'). \end{aligned}$$

The other inequality is proved similarly. □

As is evident from the proof, the number N depends only on the charts and on l_0. It is independent of η_1 and η_2, or the choice of E and T (provided of course that everything is as described before).

5. The main proposition

Using the machinery developed in Sections 2, 3 and 4 we now prove that the entropy of f is equal to the entropy of f with respect to certain partitions subordinate to W^u.

PROPOSITION 5.1. *Suppose f: $M \hookleftarrow$ is a C^2 diffeomorphism of a compact Riemannian manifold and m is an ergodic Borel probability measure on M. Let $\beta > 0$ be given. Then there is an increasing measurable partition ξ_β of the type discussed in* (3.1) *such that*

$$\beta(c + s) \geq (1 - \beta)\left[h_m(f) - h_m(f, \xi_\beta) - \beta\right].$$

Proof. Our strategy is to construct ξ_β as in (3.1) and to use it to construct η_1 and η_2 as in (3.2) with $h_m(f, \eta_2) \geq h_m(f) - \beta/3$. Calling the conditional measures associated with η_1 and η_2 $\{m_x^1\}$ and $\{m_x^2\}$ respectively, we will show that if $B^T(x, \rho) = \{y \in \eta_2(x)\colon d_x^T(x, y) < \rho\}$, then

$$\beta \cdot \liminf_{\rho \to 0} \frac{\log m_x^2 B^T(x, \rho)}{\log \rho} \geq (1 - \beta)\left[h_m(f, \eta_2) - h_m(f, \eta_1) - 2\beta/3\right]$$

for m-a.e. x. The desired conclusion follows immediately from this and Lemmas 3.1.2 and 4.1.4.

We divide the proof into 5 parts.

(A) We start by enumerating the specifications on ξ_β, η_1 and η_2. First fix $\varepsilon > 0$. We assume that $\varepsilon \leq \beta/3$, $\lambda^+/100$ and $-\lambda^-/100$. Let $\{\Phi_x, x \in \Gamma'\}$ be a system of (ε, l)-charts as described in (2.1). Using these charts, we construct an increasing measurable partition ξ_β as in the proof of Lemma 3.1.1 with S, l_0 and γ having the same meaning as in that proof. Let $N = N(l_0)$ be the constant in Lemma 4.2.3. Pick $E \subset \mathrm{S} \cap \{l \leq l_0\}$ according to (3.4) and the paragraph before Lemma 4.2.3. Let T be chosen likewise. We assume that the measure of E is small enough that $e^{-\beta\varepsilon}N^{4m(E)} < 1$. Now as in (3.2), let $\delta_0 = \min(\frac{1}{4}, \frac{1}{2}e^{-\lambda^- - \varepsilon}, \gamma/2K)$ and let $\mathscr{P}$ be a finite entropy partition adapted to $(\{\Phi_x\}, \delta_0)$. We require also that $\mathscr{P}$ refine $\{\mathrm{S}, M - \mathrm{S}\}$ and $\{E, M - E\}$ and that $h_m(f, \mathscr{P}) \geq h_m(f) - \varepsilon$. Finally we set $\eta_1 = \xi_\beta \vee \mathscr{P}^+$ and $\eta_2 = \mathscr{P}^+$. Recall that with η_1 and η_2 so constructed, $\eta_2(x)/\eta_1$ has a nice quotient structure endowed with a transverse metric d_x^T for m-a.e. x.

(B) Before proceeding with the main argument, we record some estimates derived from the results of (4.1). For $\delta > 0$, define g, g_δ, g_*: $M \to \mathbf{R}$ by

$$g(y) = m_y^1(f^{-1}\eta_2)(y),$$

$$g_\delta(y) = \frac{1}{m_y^2 B^T(y,\delta)} \int_{B^T(y,\delta)} m_z^1(f^{-1}\eta_2)(y) m_y^2(dz) \quad \text{and}$$

$$g_*(y) = \inf_{\delta>0} g_\delta(y).$$

Note that by Lemma 3.3.2 $g(y)$ is also equal to $m_y^1(f^{-1}\eta_1)(y)$. We leave it to the reader to verify the measurability of g_δ. (For fixed δ, one could check for instance that $y \mapsto \int_{B^T(y,\delta)} m_z^1(f^{-1}\eta_2)(y) m(dz)$ is measurable on E.)

We claim that $g_\delta \to g$ almost everywhere on M and that $\int -\log g_* dm < \infty$. To see this, first consider one η_2-element at a time. Fix x. Substitute $(\eta_2(x), m_x^2)$ for (X, μ) in Lemma 4.1.3, let π: $\eta_2(x) \to \mathbf{R}^{c+s}$ be the π in (3.4) and let $\alpha = (f^{-1}\eta_2)|\eta_2(x)$. Then g, g_δ and g_* as defined above agree with the corresponding functions in (4.1.3). We can therefore conclude that $g_\delta \to g$, m_x^2-a.e. and that $\int -\log g_* dm_x^2 \le H_{m_x^2}(f^{-1}\eta_2) + \log c + 1$. Integrating over M, this gives $\int -\log g_* dm \le H_m(f^{-1}\eta_2|\eta_2) + \log c + 1 < \infty$.

(C) The purpose of this step is to study the induced action of f on $f^{-1}(\eta_2(x))/\eta_1 \to \eta_2(x)/\eta_1$ with respect to the metrics $d_{f^{-1}x}^T$ and d_x^T. Consider $x \in M$. The point x will be subjected to a finite number of a.e. assumptions. Let $r_0 < r_1 < r_2 < \cdots$ be the successive times t when $f^t x \in E$ with $r_0 \le 0 < r_1$. Note that r_0 is constant on $\eta_2(x)$. For large n and $0 \le k < n$, define $a(x,k)$ as follows: If $r_j \le k < r_{j+1}$, then

$$a(x,k) = B^T\big(f^k x, e^{-\beta(n-r_j)} N^{2j}\big).$$

LEMMA 5.2. $a(x,k) \cap (f^{-1}\eta_2)(f^k x) \subset f^{-1}a(x, k+1)$.

Proof. If $k \ne r_j - 1$ for any j, then we have $fa(x,k) \cap \eta_2(f^{k+1}x) = a(x, k+1)$ automatically since $d_{f^k x}^T$ and $d_{f^{k+1}x}^T$ are defined by pulling back to E.

The case when $k = r_j - 1$ for some j reduces to the following consideration: Let $y \in E$ and let $r > 0$ be the smallest integer such that $f^r y \in E$. Let $z \in (f^{-r}\eta_2)(y)$. It suffices to show that

$$d_{f^r y}^T(f^r y, f^r z) \le N^2 e^{r\beta} d_y^T(y, z).$$

First we have $d_y'(y, z) \le N d_y^T(y, z)$. (For the definition of d_y' see (3.4).) Then for $i = 1, 2, \ldots, r$, Lemma 2.3.2 tells us that $d_{f^i y}'(f^i y, f^i z) \le e^{\beta i} d_y'(y, z)$. We pick up another factor of N when converting back to the d^T-metric at $f^r y$. $\square$

(D) We now estimate $m_x^2 B^T(x, e^{-\beta(n - r_0(x))}) = m_x^2 a(x, 0)$, which we can write as

$$m_x^2 a(x,0) = \prod_{k=0}^{p-1} \frac{m_{f^k x}^2 a(x,k)}{m_{f^{k+1}x}^2 a(x, k+1)} \cdot m_{f^p x}^2 a(x,p)$$

where $p = [n(1-\varepsilon)]$. First note that the last term ≤ 1. For each $0 \leq k < p$,

$$\frac{m_{f^k x}^2 a(x,k)}{m_{f^{k+1}x}^2 a(x,k+1)} = m_{f^k x}^2 a(x,k) \cdot \frac{m_{f^k x}^2 f^{-1}(\eta_2(f^{k+1}x))}{m_{f^k x}^2 f^{-1}(a(x,k+1))}$$

by invariance of m and uniqueness of conditional probabilities. This is

$$(*) \qquad \leq \frac{m_{f^k x}^2 a(x,k)}{m_{f^k x}^2\big((f^{-1}\eta_2)(f^k x) \cap a(x,k)\big)} \cdot m_{f^k x}^2 (f^{-1}\eta_2)(f^k x)$$

by Lemma 5.2. If g_δ is defined as in (B), then the first quotient in $(*)$ is equal to

$$\left[g_{\delta(x,n,k)}(f^k x)\right]^{-1} \quad \text{where}$$
$$\delta(x,n,k) = e^{-\beta(n - r_j(x))} N^{2j} \quad \text{and}$$
$$j = \#\{0 < i \leq k : f^i x \in E\}.$$

When $I(x) = -\log m_x^2(f^{-1}\eta_2)(x)$, the second term in $(*)$ is equal to $e^{-I(f^k x)}$. Thus

$$\log m_x^2 B^T(x, e^{-\beta(n - r_0(x))}) \leq -\sum_{k=0}^{p-1} \log g_{\delta(x,n,k)}(f^k x) - \sum_{k=0}^{p-1} I(f^k x).$$

Multiplying by $-1/n$ and taking lim inf on both sides of this inequality, we have

$$\begin{aligned} &\beta \cdot \liminf_{\rho \to 0} \frac{\log m_x^2 B^T(x,\rho)}{\log \rho} \\ &\quad = \beta \cdot \liminf_{n \to \infty} \frac{\log m_x^2 B^T(x, e^{-\beta(n - r_0(x))})}{\log e^{-\beta n}} \\ &\quad \geq \liminf_{n \to \infty} \frac{1}{n} \sum_{k=0}^{[n(1-\varepsilon)]} \log g_{\delta(x,n,k)}(f^k x) + \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{[n(1-\varepsilon)]} I(f^k x) \end{aligned}$$

where the last limit $= (1-\varepsilon) H_m(\eta_2 | f\eta_2) \geq (1-\varepsilon)(h_m(f) - \varepsilon)$. Thus Proposition 5.1 is proved if we show that

$$\limsup_n -\frac{1}{n} \sum_{k=0}^{[n(1-\varepsilon)]} \log g_{\delta(x,n,k)}(f^k x) \leq (1-\varepsilon)[h_m(f,\eta_1) + 2\varepsilon].$$

(E) We prove this last assertion. It follows from (B) that there is a measurable function $\delta(x)$ such that if $\delta \leq \delta(x)$, then $-\log g_\delta(x) \leq -\log g(x) + \varepsilon$. Also, since $\int -\log g_* < +\infty$, there is a number δ_1 such that if $A = \{\delta > \delta_1\}$ then $\int_{M-A} -\log g_* \leq \varepsilon$.

We claim that for almost every x, if n is sufficiently large, then $\delta(x, n, k) \leq \delta_1$ for all $k \leq n(1-\varepsilon)$. First there is $N(x)$ such that for $n \geq N(x)$, $\#\{0 \leq i < n\colon f^i x \in E\} \leq 2n \cdot m(E)$. If $n \geq N(x)$, then

$$\begin{aligned}\delta(x, n, k) &= e^{-\beta(n-r_j)} N^{2j}\\ &\leq e^{-\beta\varepsilon n} N^{2 \cdot 2nm(E)}.\end{aligned}$$

Since $e^{-\beta\varepsilon} N^{4m(E)} < 1$, $\delta(x, n, k)$ is less than δ_1 for n sufficiently large. Thus

$$\sum_{k=0}^{[n(1-\varepsilon)]} -\log g_{\delta(x,n,k)}(f^k x)$$

$$\leq \sum_{\substack{k=0\\ f^k x \in A}}^{[n(1-\varepsilon)]} \left(-\log g(f^k x) + \varepsilon\right) + \sum_{\substack{k=0\\ f^k x \notin A}}^{[n(1-\varepsilon)]} -\log g_*(f^k x)$$

and the lim sup we wish to estimate is bounded above by

$$(1-\varepsilon)\left[\int -\log g + \varepsilon + \int_{M-A} -\log g_*\right].$$

Recall now that $g(x) = m_x^1(f^{-1}\eta_1)(x)$, so that $\int -\log g = h_m(f, \eta_1)$. This completes the proof. $\square$

Corollary 5.3. *With the same hypotheses as in Proposition* 5.1, *if ξ is any partition constructed in the proof of Lemma* 3.1.1, *then*

$$h_m(f, \xi) = h_m(f).$$

Proof. For any β, $h_m(f, \xi) = h_m(f, \xi_\beta)$ where ξ_β is as in Proposition 5.1. Let $\beta \to 0$. $\square$

6. Proof of theorems

We fill in the gaps between the results in Section 5 and Theorems A and B as stated in Section 1.

(6.1) *Proof of Theorem* A: *the ergodic case.* We may assume that $u > 0$. (The reader can verify that Theorem A is completely trivial if $u = 0$.) Let ξ be an increasing partition subordinate to W^u, constructed as in the proof of Lemma 3.1.1. By Corollary 5.3, $H_m(\xi | f\xi) = h_m(f)$. Let $\{m_x\} = \{m_x^\xi\}$ be the conditional probabilities associated with ξ and let μ_x be the Riemannian measure on

$W^u(x)$. It remains to show that

$$H(\xi|f\xi) = \sum_i \lambda_i^+ \dim E_i \Rightarrow m_x \ll \mu_x \text{ for } m\text{-a.e. } x.$$

This proof is given in French in [L]. We recall the ideas involved for the sake of completeness.

Let $J^u = |\mathrm{Jac}(Df|E^u(x))|$. By Oseledec's Theorem, $\int \log J^u = \sum_i \lambda_i^+ \dim E_i$. *Suppose* we know that $m_x \ll \mu_x$ for almost every x. Then $dm_x = \rho d\mu_x$ almost everywhere for some function ρ. This function must satisfy $\int_{\xi(x)} \rho(y) d\mu_x(y) = 1$, and $\rho(y) J^u(f^{-1}y)/\rho(f^{-1}y)$ must be constant on $\xi(x)$ by the change of variables formula. (See [LS], Proposition 4.2.) From this we can guess that for all $y \in \xi(x)$,

$$\Delta(x, y) \overset{\text{def}}{=} \frac{\rho(y)}{\rho(x)} = \frac{\prod_{i=1}^{\infty} J^u(f^{-i}x)}{\prod_{i=1}^{\infty} J^u(f^{-i}y)}.$$

A candidate for ρ then is $\rho(y) = \Delta(x, y)/L(x)$, where $L(x) = \int_{\xi(x)} \Delta(x, y)\, d\mu_x$. Of course all this makes sense only if $\Delta(x, y)$ is uniformly bounded on $\xi(x)$.

Lemma 6.1.1. *For almost every x, $y \mapsto \log \Delta(x, y)$ is a Lipschitz function on $\Phi_x W^u_{x,1}(x)$. It follows from this that for each $x, y \mapsto \Delta(x, y)$ is uniformly bounded away from 0 and $+\infty$ on $\xi(x)$.*

Proof. This is a standard calculation relying on the Lipschitz property of the functions $z \mapsto Df_z$ and $z \mapsto E^u(z)$ (Lemma 4.2.2) and the fact that for any two points $y, y' \in W^u_{x,1}(x)$, $|\tilde{f}_x^{-n} y - \tilde{f}_x^{-n} y'| \le e^{-n(\lambda^+ - 2\varepsilon)} |y - y'|$ (Lemma 2.3.1(a)). □

So we define ρ as above and define a measure ν on M such that if $\{\nu_x\}$ are the ξ-conditional measures of ν, then $d\nu_x = \rho\, d\mu_x$ and ν coincides with μ on $\mathscr{B}_\xi$, the biggest σ-algebra containing sets that are unions of elements of ξ.

Lemma 6.1.2. $\int -\log \nu_x(f^{-1}\xi)(x)\, dm(x) = \int \log J^u dm$.

Proof. Define $q(x) = \nu_x(f^{-1}\xi)(x)$. Then

$$q(x) = \frac{\int_{(f^{-1}\xi)(x)} \Delta(x, y)\, d\mu_x(y)}{L(x)} = \frac{L(fx)}{L(x)} \cdot \frac{1}{J^u(x)}.$$

Since L is a positive finite-valued measurable function with

$$\int \log^+ (L(fx)/L(x)) \le \int \log^+ J^u < \infty,$$

it follows that $\log q$ is integrable and $\int \log q = -\int \log J^u$. (See e.g. [LS], Proposition 2.2.) $\square$

From the definition of ν, it is clear that $\nu = m$ when restricted to the σ-algebra $\mathscr{B}_\xi$. The next lemma and an induction show that they are equal on $\mathscr{B}_{f^{-n}\xi}$ for all $n \geq 0$ and hence they are equal on $\mathscr{B}$.

Lemma 6.1.3. *$\int \log J^u dm = H_m(\xi | f\xi)$ implies $\nu = m$ on $\mathscr{B}_{f^{-1}\xi}$.*

For m-a.e. x, $(f^{-1}\xi)|\xi(x)$ is a countable partition. For $y \in \xi(x)$, define

$$\frac{d\nu}{dm}\bigg|_{f^{-1}\xi}(y) = \frac{\nu_y(f^{-1}\xi)(y)}{m_y(f^{-1}\xi)(y)}.$$

Note that $(d\nu/dm)|_{f^{-1}\xi}$ is well defined almost everywhere. By the convexity of log we have

$$\int \log\left(\frac{d\nu}{dm}\bigg|_{f^{-1}\xi}\right) dm \leq \log \int \left(\frac{d\nu}{dm}\bigg|_{f^{-1}\xi}\right) dm = 0$$

with

$$\int \log\left(\frac{d\nu}{dm}\bigg|_{f^{-1}\xi}\right) dm = 0 \text{ if and only if } \frac{d\nu}{dm}\bigg|_{f^{-1}\xi} \equiv 1 \ m\text{-a.e.}$$

But we know that $\int \log(d\nu/dm)|_{f^{-1}\xi} dm = 0$, for Lemma 6.2.2 says that

$$\begin{aligned}
-\int \log \nu_x(f^{-1}\xi)(x)\, dm &= \int \log J^u dm \\
&= H_m(f^{-1}\xi | \xi) \\
&= -\int \log m_x(f^{-1}\xi)(x)\, dm.
\end{aligned}$$

Thus $\nu = m$ on $\mathscr{B}_{f^{-1}\xi}$. This completes the proof of the ergodic case of Theorem A. $\square$

Corollary 6.1.4. *Let m be an ergodic measure satisfying Pesin's formula, let ξ be as above, and let ρ be the density of m_x^ξ with respect to μ_x. Then at m-a.e. x, ρ is a strictly positive function on $\xi(x)$ satisfying*

$$\frac{\rho(y)}{\rho(x)} = \prod_{i=1}^{\infty} \frac{J^u(f^{-i}x)}{J^u(f^{-i}y)}.$$

In particular, $\log \rho$ *is Lipschitz along W^u-leaves.*

Remark. It can in fact be shown that when f is C^2, each $W^u(x)$ is a C^2 immersed submanifold (see e.g. [PS]) and that ρ is C^1 along $W^u(x)$.

(6.2) *Proof of Theorem* A. We reduce the general theorem to its ergodic case. Let $\mathfrak{g}$ be the sub-σ-algebra of $\mathscr{B}$ consisting of all invariant subsets and let ζ be a measurable partition such that $\mathscr{B}_\zeta \doteq \mathfrak{g}$. (We know that such a partition exists.) Choose a family of conditional probability measures associated with ζ. Call it $\{m_x\}$. Then there is an invariant set $N_1 \subset \Gamma'$ with $mN_1 = 1$ such that for every $x \in N_1$, m_x is invariant and ergodic.

Suppose that $h_m(f) = \int \sum_i \lambda_i^+(x) \dim E_i(x) m(dx)$. Since

$$h_m(f) = \int h_{m_x}(f) m(dx) \quad \text{and}$$

$$h_{m_x}(f) \leq \int \sum_i \lambda_i^+(y) \dim E_i(y) m_x(dy)$$

for every $x \in N_1$[Ru2], there is an invariant set $N_2 \subset N_1$ with $mN_2 = 1$ such that for every $x \in N_2$,

$$h_{m_x}(f) = \int \sum_i \lambda_i^+(y) \dim E_i(y) m_x(dy).$$

Let ξ be a measurable partition subordinate to the W^u-foliation and let $\{m_x^\xi\}$ be a family of conditional probility measures associated with ξ. We verify a couple of technical points before applying (6.1) to f: $(M, m_x) \hookleftarrow$.

First there is an invariant set $N_3 \subset N_2$ with $mN_3 = 1$ such that for every $x \in N_3$, $\xi(y) \subset W^u(y)$ and contains a neighborhood of y in $W^u(y)$ for m_x-a.e. y, i.e. ξ is indeed a partition subordinate to W_u with respect to m_x for every $x \in N_3$. More crucial is the fact that ξ refines ζ (see e.g. [LS] Proposition 2.6). This implies that there is a set $N_4 \subset N_3$ with $mN_4 = 1$ such that for every $x \in N_4$, $\{m_y^\xi\}$ is a family of conditional probability measures associated with ξ in the space $(N_4 \cap \xi(x), m_x)$.

Thus for f: $(M, m_x) \hookleftarrow$, $x \in N_4$, we can appeal to the proof of the ergodic case and conclude that m_y^ξ is absolutely continuous with respect to $\mu_{W^u(y)}$ for m_x-a.e. y. Moreover, $dm_y^\xi / d\mu_{W^u(y)}$ is as in Corollary 6.1.4.

Let $A = \{x: m_x^\xi \ll \mu_{W^u(x)}\}$. It is straightforward to verify that A is a measurable set. We have just proved that $m_x A = 1$ for m-a.e. x. Therefore $m(A) = 1$ and the proof of Theorem A is complete. □

COROLLARY 6.2. *Corollary* 6.1.4 *is true in the nonergodic case.*

(6.3) *Proof of Theorem* B. We shall prove that $\mathscr{B}^u \doteq$ the Pinsker σ-algebra. The other equality involving $\mathscr{B}^s$ is obtained by substituting f by f^{-1}. Again we first prove the theorem assuming that m is ergodic.

The case $u = 0$ is trivial, for then $\mathscr{B}^u \doteq \mathscr{B} \doteq$ the Pinsker σ-algebra. So suppose $u > 0$, and let ξ be a partition constructed in (3.1). Then $\{\mathscr{B}_{f^{-n}\xi}\}_{n \geq 0}$ is a generating family and $h_m(f) = H_m(\xi | f\xi)$. A theorem in [Ro2] tells us that for ξ with these properties, the Pinsker σ-algebra coincides with $\bigcap_{n \geq 0} \mathscr{B}_{f^n \xi}$, which in this case is $\mathscr{B}^u$.

When m is not ergodic, let $\mathfrak{g}$, ζ and $\{m_x\}$ be as in (6.2). There is a measurable partition ξ^u such that $\mathscr{B}_{\xi^u} \stackrel{\circ}{=} \mathscr{B}^u(m)$, and we let $\{m_x^u\}$ be a family of conditional probability measures associated with ξ^u.

Since $\mathfrak{g} \subset \mathscr{B}^u$, there is an invariant set N_1 with $mN_1 = 1$ such that for every $x \in N_1$ we have

i) m_x is invariant and ergodic,

ii) ξ^u is a measurable partition of $\zeta(x)$ and

$$\mathscr{B}_{\xi^u} \stackrel{\circ}{=} \mathscr{B}^u(m_x)$$

and

iii) $y \mapsto m_y^u$ is a family of conditional measures associated with ξ^u in the space $(\zeta(x), m_x)$.

Let A be a set in the Pinsker σ-algebra of f: $(M, m) \hookleftarrow$. Since $h_m(f, \{A, M - A\}) = 0$, there is an invariant set $N_2 \subset N_1$ with $mN_2 = 1$ such that $h_{m_x}(f, \{A, M - A\}) = 0$ for $x \in N_2$. For such x, our argument in the ergodic case shows that $A \cap \zeta(x)$ is in $\mathscr{B}^u$ and by ii) and iii) above, $m_y^u(A) = 0$ or 1 at m_x-a.e. y. Thus $m_x^u(A) = 0$ or 1 at m-a.e. x and therefore A is in $\mathscr{B}_{\xi^u}$. This proves that the Pinsker σ-algebra is contained in $\mathscr{B}^u$. The extension of the other containment is easy. □

Appendix: Lyapunov charts

We include here an outline of the construction of Lyapunov charts partly for the convenience of the reader and partly because we need a little more than what is usually done. (See for instance [P1].) All notation is as in Section 2. In addition, we let $\langle \cdot , \cdot \rangle$ be the usual inner product in Euclidean space, $\langle\langle \cdot , \cdot \rangle\rangle_x$ be the inner product on T_xM given by the Riemannian structure and $\| \cdot \|_x$ be its corresponding norm.

Let $\varepsilon > 0$ be given as in Section 2. We use as our starting point the following fact, namely that there is a measurable function C: $\Gamma' \to [1, \infty)$ such that:

1. For every $x \in \Gamma'$ and $n \geq 0$,

$$\|Df_x^{-n} v\|_{f_x^{-n}} \leq C(x) e^{-(\lambda^+ - \varepsilon/2)n} \|v\|_x \quad \text{for all } v \in E^u(x),$$

$$\|Df_x^{n} v\|_{f_x^{n}} \leq C(x) e^{(\lambda^- + \varepsilon/2)n} \|v\|_x \quad \text{for all } v \in E^s(x),$$

$$\|Df_x^{\pm n} v\|_{f^{\pm n} x} \leq C(x) e^{(\varepsilon/2)n} \|v\|_x \quad \text{for all } v \in E^c(x);$$

2. $|\sin(E^u(x), E^c(x))|$, $|\sin(E^c(x), E^s(x))|$, $|\sin(E^s(x), E^u(x))| \geq C(x)^{-1}$;
3. $C(f^{\pm}x) \leq e^{(1/3)\varepsilon}C(x)$.

A standard technique for obtaining good estimates on Df after *one* iteration is to introduce a new inner product $\langle\langle\ \cdot\ ,\cdot\rangle\rangle'_x$ on T_xM for every $x \in \Gamma'$. First we let

$$\langle\langle u, v\rangle\rangle'_x = \begin{cases} \displaystyle\sum_{n=0}^{\infty} \frac{\langle\langle Df_x^{-n}u, Df_x^{-n}v\rangle\rangle_{f^{-n}x}}{e^{-2n(\lambda^+ - \varepsilon)}} & \text{for } u, v \in E^u(x) \\ \displaystyle\sum_{n=-\infty}^{\infty} \frac{\langle\langle Df_x^{n}u, Df_x^{n}v\rangle\rangle_{f^{n}x}}{e^{2|n|\varepsilon}} & \text{for } u, v \in E^c(x) \\ \displaystyle\sum_{n=0}^{\infty} \frac{\langle\langle Df_x^{n}u, Df_x^{n}v\rangle\rangle_{f^{n}x}}{e^{2n(\lambda^- + \varepsilon)}} & \text{for } u, v \in E^s(x). \end{cases}$$

Then we extend $\langle\langle\ \cdot\ ,\cdot\rangle\rangle'_x$ to all of T_xM by demanding that the subspaces $E^u(x)$, $E^c(x)$ and $E^s(x)$ be mutually orthogonal with respect to $\langle\langle\ \cdot\ ,\cdot\rangle\rangle'_x$.

Recall that our objective here is to define a measurable function $l\colon \Gamma' \to [1, \infty)$ and a family of maps $\{\Phi_x\colon R(l(x)^{-1}) \to M, x \in \Gamma'\}$ so that these coordinate charts and their connecting maps have properties i)–iv) as stated in (2.1). Let $L_x\colon T_xM \to \mathbf{R}^{u+c+s}$ be a linear map taking $E^u(x)$, $E^c(x)$ and $E^s(x)$ onto $\mathbf{R}^u \times \{0\} \times \{0\}$, $\{0\} \times \mathbf{R}^c \times \{0\}$ and $\{0\} \times \{0\} \times \mathbf{R}^s$ respectively and satisfying

$$\langle L_xu, L_xv\rangle = \langle\langle u, v\rangle\rangle'_x$$

for every $u, v \in T_xM$. Setting

$$\Phi_x = \exp_x \circ L_x^{-1},$$

one immediately verifies i) and ii) in (2.1).

Next we want some bounds on $\|v\|'_x/\|v\|_x$ for $v \neq 0 \in T_xM$, where $\|\cdot\|'_x$ is the norm derived from $\langle\langle\ \cdot\ ,\cdot\rangle\rangle'_x$. First we consider $v \in E^u(x)$ or $E^c(x)$ or $E^s(x)$. Obviously $\|v\|'_x \geq \|v\|$. A direct calculation using the properties of the function $C(x)$ shows that

$$\|v\|'_x \leq C_0C(x)\|v\|_x$$

where $C_0 = (2\sum_{n=0}^{\infty}e^{-\varepsilon n})^{1/2}$. For arbitrary $v \in T_xM$, write $v = v^u + v^c + v^s$ respecting the decomposition $E^u(x) \oplus E^c(x) \oplus E^s(x)$. It is easy to check that $\|v\|_x \leq 3\|v\|'_x$. Observe that $\|v\| \geq \|v^u\| \cdot |\sin\theta_1| \cdot |\sin\theta_2|$ where θ_1 is the angle between $v^u + v^c$ and v^s and θ_2 is the angle between v^u and v^c. Since $|\sin\theta_1|$,

$|\sin\theta_2| \geq C(x)^{-1}$, we have

$$\|v\|'_x \leq C_0 C(x)\left[\|v^u\|_x + \|v^c\|_x + \|v^s\|_x\right] \leq 3C_0 C(x)^3 \|v\|_x.$$

Finally we are in a position to show that properties iii) and iv) are satisfied if we let $l(x) = C \cdot C(x)^3$ where C is a constant the magnitude of which will be obvious from the next discussion:

There is number $\delta > 0$ such that for all $x \in \Gamma'$, $\exp_x$ restricted to $\{\|v\| \leq \delta\}$ is a diffeomorphism with $\|D\exp_x\|$, $\|D\exp_x^{-1}\| \leq 2$. If C is large enough, we are assured that $L_x^{-1}R(l(x)^{-1}) \subset \{\|v\| \leq \delta\}$ and that property iv) holds with $K = 6$. Since

$$\tilde{f}_x = L_{fx} \circ \exp_{fx}^{-1} \circ f \circ \exp_x \circ L_x^{-1},$$

and the second derivatives of exp, $\exp^{-1}$ and f are uniformly bounded in x, the Lipschitz size of $D\tilde{f}_x$ is essentially determined by the norm of L_{fx}. Thus for C sufficiently large we have

$$L(D\tilde{f}_x) \leq l(x)\varepsilon.$$

Also,

$$|D\tilde{f}_x(z) - D\tilde{f}_x(0)| \leq l(x)\varepsilon|z| \leq \varepsilon$$

for $z \in R(l(x)^{-1})$. Similar considerations for $\tilde{f}_x^{-1}$ guarantee the properties listed in iii).

Remarks. 1) The above construction requires only that f be $C^{1+\alpha}$ for some $\alpha > 0$. Needless to say, with this hypothesis property iii) in (2.1) has to be modified accordingly.

2) The reader can verify easily that if $T_xM = E_1(x) \oplus \cdots \oplus E_{r(x)}(x)$ is the decomposition into subspaces corresponding to exponents $\lambda_1(x), \ldots, \lambda_{r(x)}(x)$, then the same trick used for $E^c(x)$ above can be performed on each $E_i(x)$ separately to obtain a norm $\|\cdot\|''_x$, with the property that for each $v \in E_i(x)$,

$$e^{(\lambda_i - \varepsilon)}\|v\|''_x \leq \|Df_x v\|''_{fx} \leq e^{(\lambda_i + \varepsilon)}\|v\|''_x.$$

Charts with these properties are extremely useful in Part II.

3) If the measure m is not ergodic, the construction described in this appendix can be carried out on invariant sets of the form

$$\Gamma(\varepsilon, \lambda^+, \lambda^-) = \left\{x \in \Gamma' : \min_{\lambda_i(x) > 0} \lambda_i(x) \geq \lambda^+, \max_{\lambda_i(x) < 0} \lambda_i(x) \leq \lambda^-\right\}$$

where $\lambda^+, -\lambda^- \geq 100\varepsilon$ and $m\Gamma(\varepsilon, \lambda^+, \lambda^-) > 0$.

Université de Paris VI, Paris, France
Michigan State University, East Lansing, Michigan

References

[A] D. V. Anosov, Geodesic flow on closed Riemannian manifolds with negative curvature, Proc. Steklov Inst. of Maths. **90** (1967).

[B] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeormorphisms*, Springer Lecture Notes in Math. **470** (1975).

[FHY] A. Fathi, M. Herman and J-C. Yoccoz, A proof of Pesin's stable manifold theorem, in *Geometric Dynamics*, Springer Lecture Notes in Math. **1007** (1983), 117–215.

[G] M. de Guzmán, *Differentiation of Integrals in* $\mathbf{R}^n$, Springer Lecture Notes in Math. **481** (1975), 2.

[HP] M. W. Hirsch and C. Pugh, Stable manifolds and hyperbolic sets. Proc. Sym. in Pure Math. **14**, A.M.S., Providence, RI (1970).

[Ka] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, Publ. Math. I.H.E.S. **51** (1980), 137–174.

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

[L] F. Ledrappier, Propriétés ergodiques des mesures de Sinai, Publ. Math. I.H.E.S. **59** (1984), 163–188.

[LS] F. Ledrappier and J.-M. Strelcyn, A proof of the estimation from below in Pesin's entropy formula, Ergod. Th. & Dynam. Sys. **2** (1982), 203–219.

[LY] F. Ledrappier and L.-S. Young The metric entropy of diffeomorphisms, Bull. A.M.S. **11** (1984), 343–346.

[M] R. Mañé, A proof of Pesin's formula, Ergod. Th. and Dynam. Sys. **1** (1981), 95–102.

[O] V. I. Oseledec, Multiplicative ergodic theorem. Lyapunov Characteristics numbers for dynamical systems, Trans. Mosc. Math. Soc. **19** (1968), 197–221.

[P1] Ya. B. Pesin, Families of invariant manifolds corresponding to non-zero characteristics exponents, Math. of the USSR, Izvestija **10** (1978), 1261–1305.

[P2] ______, Characteristic Lyapunov exponents and smooth ergodic theory, Russ. Math. Surveys **32**:4 (1977), 55–114.

[PS] C. Pugh and M. Shub, *Ergodic Attractors*, in preparation.

[Ro1] V. A. Rohlin, On the fundamental ideas of measure theory, A.M.S. Transl. (1) **10** (1962), 1–52.

[Ro2] V. A. Rohlin, Lectures on the theory of entropy of transformations with invariant measures, Russ. Math. Surveys **22**:5 (1967), 1–54.

[Ru1] D. Ruelle, A measure associated with Axiom A attractors, Amer. J. Math. **98** (1976), 619–654.

[Ru2] ______, An inequality for the entropy of differentiable maps, Bol. Soc. Bras. Math. **9** (1978), 83–87.

[Ru3] ______, Ergodic theory of differentiable dynamical systems, Publ. Math. I.H.E.S. **50** (1979), 27–58.

[S1] Ya. G. Sinai, Classical systems with countable Lebesgue spectrum II, A.M.S. Transl. **68** (1968), 34–88.

[S2] ______, Markov partitions and C-diffeomorphisms, Func. Anal. and its Appl. **2** (1968), 64–89.

[S3] ______, Gibbs measures in ergodic theory, Russ. Math. Surveys **27**:4 (1972), 21–69.