SMOOTH ERGODIC THEORY

AMIE WILKINSON

Contents

Glossary

Conservative, Dissipative: Conservative dynamical systems (on a compact phase space) are those that preserve a finite measure equivalent to volume. Hamiltonian dynamical systems are important examples of conservative systems. Systems that are not conservative are called dissipative. Finding physically meaningful invariant measures for dissipative maps is a central object of study in smooth ergodic theory.

Distortion estimate: A key technique in smooth ergodic theory, a distortion estimate for a smooth map f gives a bound on the variation of the jacobian of f^n in a given region, for n arbirtarily large. The jacobian of a smooth map at a point x is the absolute value of the determinant of derivative at x, measured in a fixed Riemannian metric. The jacobian measures the distortion of volume under f in that metric.

Hopf Argument: A technique developed by Eberhard Hopf for proving that a conservative diffeomorphism or flow is ergodic. The argument relies on the Ergodic Theorem for invertible transformations, the density of continuous functions among integrable functions, and the existence of stable and unstable foliations for the system. The argument has been used, with various modifications, to establish ergodicity for hyperbolic, partially hyperbolic and nonuniformly hyperbolic systems.

Hyperbolic: A compact invariant set $\Lambda \subset M$ for a diffeomorphism $f\colon M \to M$ is hyperbolic if, at every point in Λ, the tangent space splits into two subspaces, one that is uniformly contracted by the derivative of f, and another that is uniformly expanded. Expanding maps and Anosov diffeomorphisms are examples of globally hyperbolic maps. Hyperbolic diffeomorphisms and flows are the archetypical smooth systems displaying chaotic behavior, and their dynamical properties are well-understood. Nonuniform hyperbolicity and partial hyeprbolicity are two generalizations of hyperbolicity that encompass a broader class of systems and display many of the chaotic features of hyperbolic systems.

Sinai-Ruelle-Bowen (SRB) measure: The concept of SRB measure is a rigorous formulation of what it means for an invariant measure to be "physically meaningful." An SRB measure attracts a large set of orbits into its support, and its statistical features are reflected in the behavior of these attracted orbits.

Definition and Importance of the Subject

Smooth ergodic theory is the study of the statistical and geometric properties of measures invariant under a smooth transformation or flow. The study of smooth ergodic theory is as old as the study of abstract ergodic theory, having its origins in Bolzmann's Ergodic Hypothesis in the late 19th Century. As a response to Boltzmann's hypothesis, which was formulated in the context of Hamiltonian Mechanics, Birkhoff and von Neumann defined ergodicity in the 1930's and proved their foundational ergodic theorems. The study of ergodic properties of smooth systems saw an advance in the work of Hadamard and E. Hopf in the 1930's their study of geodesic flows for negatively curved surfaces. Beginning in the 1950's, Kolmogorov, Arnold and Moser developed a perturbative theory producing obstructions to ergodicity in Hamiltonian systems, known as Kolmogorov-Arnold-Moser (KAM) Theory. Beginning in the 1960's with the work of Anosov and Sinai on hyperbolic systems, the study of smooth ergodic theory has seen intense activity. This activity continues today, as the ergodic properties of systems displaying weak forms of hyperbolicity are further understood, and KAM theory is applied in increasingly broader contexts.

1. Introduction

This entry focuses on the basic arguments and principles in smooth ergodic theory, illustrating with simple and straightforward examples. The classic texts [1, 2] are a good supplement.

The discussion here sidesteps the topic of Kolmogorov-Arnold-Moser (KAM) theory, which has played an important role in the development of smooth ergodic theory; see the entry in this volume. For reasons of space, detailed discussion of several active areas in smooth ergodic theory is omitted, including: higher mixing properties (Kolmogorov, Bernoulli, etc.), finer statistical properties (fast decay of correlations, Central Limit Theorem, large deviations), smooth thermodynamic formalism (transfer operators, pressure, dynamical zeta functions, etc.), the smooth ergodic theory of random dynamical systems, as well as any mention of infinite invariant measures. The text [3] covers many of these topics, and the texts [4, 5, 6] treat random smooth ergodic theory in depth. An excellent discussion of many of the recent developments in the field of smooth ergodic theory is [7].

This entry assumes knowledge of the basic concepts in ergodic theory and of basic differential topology. The texts [8] and [9] contain the necessary background.

2. The volume class

For simplicity, assume that M is a compact, boundaryless C^∞ Riemannian manifold, and that $f\colon M \to M$ is an orientation-preserving, C^1 map satisfying $m(D_xf) > 0$, for all $x \in M$, where

$$m(D_xf) = \inf_{v\in T_xM, \|v\|=1} \|D_xf(v)\|.$$

If f is a diffeomorphism, then this assumption is automatically satisfied, since in that case $m(D_xf) = \|D_{f(x)}f^{-1}\|^{-1} > 0$. For non-invertible maps, this assumption is essential in much of the following discussion. The Inverse Function Theorem implies that any map f satisfying these hypotheses is a covering map of positive degree $d \geq 1$.

These assumptions will avoid the issues of infinite measures and the behavior of f near critical points and singularities of the derivative. For most results discussed in this entry, this assumption is not too restrictive. The existence of critical points and other singularities is, however, a complication that cannot be avoided in many important applications. The ergodic-theoretic analysis of such examples can be considerably more involved, but contains many of the elements discussed in this entry. The discussion in Section 8 indicates how some of these additional technicalities arise and can be overcome. For simplicity, the discussion here is confined almost exclusively to discrete time evolution. Many, though not all, of the the results mentioned here carry over to flows and semiflows using, for example, a cross-section construction (see [2], Chapter 1).

Every smooth map $f\colon M \to M$ satisfying these hypotheses preserves a natural measure *class*, the measure class of a finite, smooth Riemannian volume on M. Fix such a volume ν on M. Then there exists a continuous, positive *jacobian* function $x \mapsto \mathrm{jac}_xf$ on M, with the property that for every sufficiently small ball $B \subset M$, and every measurable set $A \subset B$ one has:

$$\nu(f(A)) = \int_B \mathrm{jac}_xf\, d\nu(x).$$

The jacobian of f at x is none other than the absolute value of the determinant of the derivative D_xf (measured in the given Riemannian metric). To see that the measure class of ν is preserved by f, observe that the Radon-Nikodym derivative $\frac{df_*\nu}{d\nu}(x)$ at x is equal to $\sum_{y\in f^{-1}(x)}(\mathrm{jac}_yf)^{-1} > 0$. Hence $f_*\nu$ is equivalent to ν, and f preserves the measure class of ν.

In many contexts, the map f has a natural *invariant* measure in the measure class of volume. In this case, f is said to be *conservative.* One setting in which a natural invariant smooth measure appears is Hamiltonian dynamics. Any solution to Hamilton's equations preserves a smooth volume called the *Liouville measure.* Furthermore, along the invariant, constant energy hypermanifolds of a Hamiltonian flow, the Liouville measure decomposes smoothly into invariant measures, each of which is equivalent to the induced Riemannian volume. In this way, many systems of physical or geometric origin, such as billiards, geodesic flows, hard sphere gases, and evolution of the n-body problem give rise to smooth conservative dynamical systems. See the entry on Dynamics of Hamiltonian Systems.

Note that even though f preserves a smooth measure class, it might not preserve any measure in that measure class. Consider, for example, a diffeomorphism $f\colon S^1 \to S^1$ of the circle with exactly two fixed points, p and q, $f'(p) > 1 > f'(q) >$

0. Let μ be an f-invariant probability measure. Let I be a neighborhood of p. Then $\bigcap_{n=1}^{\infty} f^{-n}(I) = \{p\}$, but on the other hand, $\mu(f^{-n}(I)) = \mu(I) > 0$, for all n. This implies that $\mu(\{p\}) > 0$, and so μ does not lie in the measure class of volume. This is an example of a dissipative map. A map f is called *dissipative* if every f-invariant measure with full support has a singular part with respect to volume. As was just seen, if a diffeomorphism f has a periodic sink, then f is dissipative; more generally, if a diffeomorphism f has a periodic point p of period k such that $\mathrm{jac}_p f^k \neq 1$, then f is dissipative.

3. The fundamental questions

For a given smooth map $f\colon M \to M$, there are the following fundamental questions.

(1) Is f conservative? That is, does there exist an invariant measure in the class of volume? If so, is it unique?
(2) When f is conservative, what are its statistical properties? Is it ergodic, mixing, a K-system, Bernoulli, etc? Does it obey a Central Limit Theorem, fast decay of correlations, large deviations estimates, etc?
(3) If f is dissipative, does there exist an invariant measure, not in the class of volume, but (in some sense) natural with respect to volume? What are the statistical properties of such a measure, if it exists?

There are several plausible ways to "answer" these questions. One might fix a given map f of interest and ask these questions for that specific f. What tends to happen in the analysis of a single map f is that either:

- the question can be answered using "soft" methods, and so the answer applies not only to f but to perturbations of f, or even to *generic* or *typical* f inside a class of maps; or
- the proof requires "hard" analysis or precise asymptotic information and cannot possibly be answered for a specific f, but can be answered for a large set of f_t in a typical (or given) parametrized family $\{f_t\}_{t\in(-1,1)}$ of smooth maps containing $f = f_0$.

Both types of results appear in the discussion that follows.

4. Lebesgue measure and local properties of volume

Locally, any measure in the measure class of volume is, after a smooth change of coordinates, equivalent to Lebesgue measure in $\mathbb{R}^n$. In fact, more is true: Moser's Theorem implies that locally any Riemannian volume is, after a smooth change of coordinates, *equal* to Lebesgue measure in $\mathbb{R}^n$. Hence to study many of the local properties of volume, it suffices to study the same properties for Lebesgue measure.

One of the basic properties of Lebesgue measure is that every set of positive Lebesgue measure can be approximated arbitrarily well in measure from the outside by an open set, and from the inside by a compact set. A consequence of this property, of fundamental importance in smooth ergodic theory, is the following statement.

Fundamental Principle #1: *Two disjoint, positive Lebesgue measure sets cannot mix together uniformly at all scales.*

As an illustration of this principle, consider the following elementary exercise in measure theory. First, some notation. If ν is a measure and A and B are ν-measurable sets with $\nu(B) > 0$, the *density of A in B* is defined by:

$$\nu(A\colon B) = \frac{\nu(A \cap B)}{\nu(B)}.$$

Proposition 1. *Let $\mathcal{P}_1, \mathcal{P}_2, \ldots$ be sequence of (mod 0) finite partitions of the circle S^1 into open intervals, with the properties: a) any element of $\mathcal{P}_n$ is a (mod 0) union of elements of $\mathcal{P}_{n+1}$, and b) the maximum diameter of elements of $\mathcal{P}_n$ tends to 0 as $n \to \infty$.*

Let A be any set of positive Lebesgue measure in S^1. Then there exists a sequence of intervals $I_1, I_2, \ldots$, with $I_n \in \mathcal{P}_n$ such that:

$$\lim_{n\to\infty} \lambda(A\colon I_n) = 1.$$

Proof. Assume that Lebesgue measure has been normalized so that $\lambda(S^1) = 1$. Fix a (mod 0) cover of $S^1 \setminus A$ by pairwise disjoint elements $\{J_i\}$ of the union $\bigcup_{n=1}^{\infty} \mathcal{P}_n$ with the properties:

$$\lambda(J_1) \geq \lambda(J_2) \geq \cdots, \text{and}$$

$$\lambda(\bigcup_{i=1}^{\infty} J_i) = \sum_{i=1}^{\infty} \lambda(J_i) < 1.$$

For $n \in \mathbb{N}$, let U_n be the union of all the intervals J_i that are contained in $\mathcal{P}_n$, and let $V_n = \bigcup_{i=1}^{n} U_n$. This defines an increasing sequence of natural numbers $i_1 = 1 < i_2 < i_3 < \cdots$ such that $U_n = \bigcup_{i=i_n}^{i_{n+1}-1} J_n$ and $V_n = \bigcup_{i=1}^{i_{n+1}-1} J_n$.

For each n, the interval I_n will be chosen to be an element $\mathcal{P}_n$, disjoint from V_n, in which the density of $\bigcup_{i=n+1}^{\infty} U_i$ is very small (approaching 0 as $n \to \infty$). Since $(S^1 \setminus A) \cap I_n$ is contained in $\bigcup_{i=n+1}^{\infty} U_i$, this choice of I_n will ensure that the density of A in I_n is large (approaching 1 as $n \to \infty$).

To make this choice of I_n, note first that the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of $S^1 \setminus V_n$ is:

$$\frac{\lambda(\bigcup_{i=n+1}^{\infty} U_i)}{\lambda(S^1 \setminus V_n)} = \frac{\sum_{i=i_{n+1}}^{\infty} \lambda(J_i)}{1 - \sum_{i=1}^{i_{n+1}-1} \lambda(J_i)} = a_n.$$

Note that, since $\sum_{i=1}^{\infty} \lambda(J_i) < 1$, one has $a_n \to 0$ as $n \to \infty$. Since the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of $S^1 \setminus V_n$ is at most a_n, there is an interval I_n in $\mathcal{P}_n$, disjoint from V_n, such that the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of I_n is at most a_n. Then

$$\lim_{n\to\infty} \lambda(A\colon I_n) \geq \lim_{n\to\infty} 1 - a_n = 1.$$

$\square$

In smooth ergodic theory, it is often useful to use a variation on Proposition 1 (generally, in higher dimensions) in which the partitions $\mathcal{P}_n$ are nested, dynamically-defined partitions. A simple application of this method can be used to prove that the doubling map on the circle is ergodic with respect to Lebesgue measure, which is done in Section 5.

Notice that this proposition does not claim that the intervals I_n are nested. If one imposes stronger conditions on the partitions $\mathcal{P}_n$, then one can draw stronger conclusions.

A very useful theorem in this respect is the Lebesgue Density Theorem. A point $x \in M$ is a *Lebesgue density point* of a measurable set $X \subseteq M$ if

$$\lim_{r\to 0} m(X\colon B_r(x)) = 1,$$

where $B_r(x)$ is the Riemannian ball of radius r centered at x. Notice that the notion of Lebesgue density point depends only on the smooth structure of M, because any two Riemannian metrics have the same Lebesgue density points. The Lebesgue Density Theorem states that if A is a measurable set and $\widehat{A}$ is the set of Lebesgue density points of A, then $m(A \,\Delta\, \widehat{A}) = 0$.

5. Ergodicity of the basic examples

This section contains proofs of the ergodicity of two basic examples of conservative smooth maps: irrational rotations on the circle and the doubling map on the circle. See the entry Ergodic Theory: Basic Examples and Constructions for a more detailed description of these maps. These proofs serve as an elementary illustration of some of the fundamental techniques and principles in smooth ergodic theory.

Rotations on the circle. Denote by S^1 the circle $\mathbb{R}/\mathbb{Z}$, which is an additive group, and by λ normalized Lebesgue-Haar measure on S^1. Fix a real number $\alpha \in \mathbb{R}$. The rotation $R_\alpha\colon S^1 \to S^1$ is the translation defined by $R_\alpha(x) = x + \alpha$. Since translations preserve Lebesgue-Haar measure, the map R_α is conservative. Note that R_α is a diffeomorphism and an isometry with respect to the canonical flat metric (length) on S^1.

Proposition 2. *If $\alpha \notin \mathbb{Q}$, then the rotation $R_\alpha\colon S^1 \to S^1$ is ergodic with respect to Lebesgue measure.*

Proof. Let A be an R_α-invariant set in S^1, and suppose that $0 < \lambda(A) < 1$. Denote by A^c the complement of A in S^1. Fix $\varepsilon > 0$. Proposition 1 implies that there exists an interval $I \subset S^1$ such that the density of A in I is large: $\lambda(A\colon I) > 1 - \varepsilon$. Similarly, one may choose an interval J such that $\lambda(A^c\colon J) > 1-\varepsilon$. Without loss of generality, one may choose I and J to have the same length. Since α is irrational, R_α has a dense orbit, which meets the interval I. Since R_α is an isometry, this impies that there is an integer n such that $\lambda(R_\alpha^n(I) \,\Delta\, J) < \varepsilon\lambda(I)$. Since $\lambda(I) = \lambda(J)$, this readily implies that $|\lambda(A\colon R_\alpha^n(I)) - \lambda(A\colon J)| < \varepsilon$. Also, since A is invariant, and R_α is invertible and preserves measure, one has:

$$\lambda(A\colon R_\alpha^n(I)) = \lambda(R_\alpha^n(A)\colon R_\alpha^n(I)) = \lambda(A\colon I) > 1 - \varepsilon.$$

But for ε sufficiently small, this contradicts the facts that $\lambda(A\colon J) = 1-\lambda(A^c\colon J) < \varepsilon$ and $|\lambda(A\colon R_\alpha^n(I)) - \lambda(A\colon J)| < \varepsilon$. $\square$

Note that this is not a proof of the strongest possible statement about R_α (namely, minimality and unique ergodicity). The point here is to show how "soft" arguments are often sufficient to establish ergodicity; this proof uses no more about R_α than the fact that it is a transitive isometry. Hence the same argument shows:

Theorem 5.1. *Let $f\colon M \to M$ be a transitive isometry of a Riemannian manifold M. Then f is ergodic with respect to Riemannian volume.*

One can isolate from this proof a useful principle:

Fundamental Principle #2: *Isometries preserve Lebesgue density at all scales, for arbitrarily many iterates.*

This principle implies, for example, that a smooth action by a compact Lie group on M is ergodic along typical (nonsingular) orbits. This principle is also useful in studying area-preserving flows on surfaces and, in a refined form, unipotent flows on homogeneous spaces. In the case of surface flows, ergodicity questions can be reduced to a study of interval exchange transformations. See the entries Basic Examples and Billiards in this volume for a detailed discussion of interval exchange transformations and flows on surfaces. The entry on Ergodic Theory of Group Actions contains detailed information on unipotent flows.

Doubling map on the circle. Let $T_2\colon S^1 \to S^1$ be the doubling map defined by $T_2(x) = 2x$. Then T_2 is a degree-2 covering map and endomorphism of S^1 with constant jacobian $\mathrm{jac}_x T_2 \equiv 2$. Since $\frac{d(T_2)_*\lambda}{d\lambda} = \frac{1}{2} + \frac{1}{2} = 1$, T_2 preserves Lebesgue-Haar measure. The doubling map is the simplest example of a hyperbolic dynamical system, a topic treated in depth in the next section.

As with the previous example, the focus here is on the property of ergodicity. It is again possible to prove much stronger results about T_2, such as Bernoullicity, by other methods. Instead, here is a soft proof of ergodicity that will generalize readily to other contexts.

Proposition 3. *The doubling map $T_2\colon S^1 \to S^1$ is ergodic with respect to Lebesgue measure.*

Proof. Let A be a T_2-invariant set in S^1 with $\lambda(A) > 0$. Let $p \in S^1$ be the fixed point of T_2, so that $T_2(p) = p$. For each $n \in \mathbb{N}$, the preimages of p under T_2^{-n} define a (mod 0) partition $\mathcal{P}_n$ into 2^n open intervals of length 2^{-n}; the elements of $\mathcal{P}_n$ are the connected components of $S^1 \setminus T_2^{-n}(\{p\})$. Note that the sequence of partitions $\mathcal{P}_1, \mathcal{P}_2, \ldots$ is nested, in the sense of Proposition 1. Restricted to any interval $J \in \mathcal{P}_n$, the map T_2^n is a diffeomorphism onto $S^1 \setminus \{p\}$ with constant jacobian $\mathrm{jac}_x(T_2^n) = (T_2^n)'(x) = 2^n$.

Since A is invariant, it follows that $T_2^{-n}(A) = A$. Fix $\varepsilon > 0$. Proposition 1 implies that there exists an $n \in \mathbb{N}$ and an interval $J \in \mathcal{P}_n$ such that $\lambda(A\colon J) > 1 - \varepsilon$. Note that $T_2^n(A \cap J) \subset A$. But then

$$\begin{aligned}
\lambda(A) &\geq \lambda(T_2^n(A \cap J)) \\
&= \int_{A\cap J} \mathrm{jac}_x(T_2^n)\, d\lambda(x) \\
&= 2^n \lambda(A \cap J) \\
&= 2^n \lambda(A\colon J)\lambda(J) \\
&> 2^n (1-\varepsilon)\lambda(J) = 1 - \varepsilon.
\end{aligned}$$

Since ε was arbitrary, one obtains that $\lambda(A) = 1$. □

In this proof, the facts that the intervals in $\mathcal{P}_n$ have constant length 2^{-n} and that the jacobian of T_2^n restricted to such an interval is constant and equal to 2^n are not essential. The key fact really used in this proof is the assertion that the ratio:

$$\frac{\lambda(T_2^n(A \cap J) \colon T_2^n(J))}{\lambda(A \colon J)}$$

is bounded, *independently of* n. In this case, the ratio is 1 for all n because T_2 has constant jacobian.

It is tempting to try to extend this proof to other expanding maps on the circle, for example, a C^1, λ-preserving map $f \colon S^1 \to S^1$ with $d_{C^1}(f, T_2)$ small. Many of the aspects of this proof carry through *mutatis mutandis* for such an f, save for one. A C^1-small perturbation of T_2 will in general no longer have constant jacobian, and the *variation* of the jacobian of f^n on a small interval can be (and often is) unbounded. The reason for this unboundedness is a lack of control of the modulus of continuity of f'. Hence this argument can fail for C^1 perturbations of T_2. On the other hand, the argument still works for C^2 perturbations of T_2, even when the jacobian is not constant.

The principle behind this fact can be loosely summarized:

Fundamental Principle #3: *On controlled scales, iterates of* C^2 *expanding maps distort Lebesgue density in a controlled way.*

This principle requires further explanation and justification, which will come in the following section. The C^2 hypothesis in this principle accounts for the fact that almost all results in smooth ergodic theory assume a C^2 hypothesis (or something slightly weaker).

6. Hyperbolic systems

One of the most developed areas of smooth ergodic theory is in the study of hyperbolic maps and attractors. This section defines hyperbolic maps and attractors, provides examples, and investigates their ergodic properties. See [10, 11] and the entry Hyperbolic Dynamical Systems in this series for a thorough discussion of the topological and smooth properties of hyperbolic systems.

A *hyperbolic structure* on a compact f-invariant set $\Lambda \subset M$ is given by a Df-invariant splitting $T_\Lambda M = E^u \oplus E^s$ of the tangent bundle over Λ and constants $C, \mu > 1$ such that, for every $x \in \Lambda$ and $n \in \mathbb{N}$:

$$v \in E^u(x) \quad \Longrightarrow \quad \|D_x f^n(v)\| \geq C^{-1}\mu^n \|v\|, \quad \text{and}$$
$$v \in E^s(x) \quad \Longrightarrow \quad \|D_x f^n(v)\| \leq C\mu^{-n} \|v\|.$$

A hyperbolic attractor for a map $f \colon M \to M$ is given by an open set $U \subset M$ such that: $f(U) \subset \overline{U}$, and such that the set $\Lambda = \bigcap_{n \geq 0} f^n(U)$ carries a hyperbolic structure. The set Λ is called the *attractor*, and U is an *attracting region*. A map $f \colon M \to M$ is *hyperbolic* if M decomposes (mod 0) into a finite union of attracting regions for hyperbolic attractors. Typically one assumes as well that the restriction of f to each attractor Λ_i is topologically transitive.

Every point p in a hyperbolic set Λ has smooth *stable manifold* $\mathcal{W}^s(p)$ and *unstable manifold* $\mathcal{W}^u(p)$, tangent, respectively, to the subspaces $E^s(p)$ and $E^u(p)$. The set $\mathcal{W}^s(p)$ is precisely the set of $q \in M$ such that $d(f^n(p), f^n(q))$ tends to 0 as $n \to \infty$, and it follows that $f(\mathcal{W}^s(p)) = \mathcal{W}^s(f(p))$. When f is a diffeomorphism, the unstable manifold $\mathcal{W}^u(p)$ is uniquely defined and is the stable manifold of f^{-1}. When f is not invertible, local unstable manifolds exist, but generally are not unique. If Λ is a transitive hyperbolic attractor, then every unstable manifold of every point $p \in \Lambda$ is dense in Λ.

Examples of hyperbolic maps and attractors

Expanding maps. The previous section mentioned briefly the C^r perturbations of the doubling map T_2. Such perturbations (as well as T_2 itself) are examples of *expanding maps.* A map $f\colon M \to M$ is *expanding* if there exist constants $\mu > 1$ and $C > 0$ such that, for every $x \in M$, and every nonzero vector $v \in T_xM$:

$$\|D_x f^n(v)\| \geq C\mu^n \|v\|,$$

with respect to some (any) Riemannian metric on M. An expanding map is clearly hyperbolic, with $U = M$, E^s the trivial bundle, and $E^u = TM$. Any disk in M is a local unstable manifold for f.

Anosov diffeomorphisms. A diffeomorphism $f\colon M \to M$ is called *Anosov* if the tangent bundle splits as a direct sum $TM = E^u \oplus E^s$ of two Df-invariant subbundles, such that E^u is uniformly expanded and E^s is uniformly contracted by Df. Similarly, a flow $\varphi_t\colon M \to M$ is called Anosov if the tangent bundle splits as a direct sum $TM = E^u \oplus E^0 \oplus E^s$ of three $D\varphi_t$-invariant subbundles, such that E^0 is generated by $\dot{\varphi}$, E^u is uniformly expanded and E^s is uniformly contracted by $D\varphi_t$. Like expanding maps, an Anosov diffeomorphism is an Anosov attractor with $\Lambda = U = M$.

A simple example of a conservative Anosov diffeomorphism is a hyperbolic linear automorphism of the torus. Any matrix $A \in SL(n, \mathbb{Z})$ induces an automorphism of $\mathbb{R}^n$ preserving the integer lattice $\mathbb{Z}^n$, and so descends to an automorphism $f_A\colon T^n \to T^n$ of the n-torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$. Since the determinant of A is 1, the diffeomorphism f_A preserves Lebesgue-Haar measure on T^n. In the case where none of the eigenvalues of A have modulus 1, the resulting diffeomorphism f_A is Anosov. The stable bundle E^s at $x \in T^n$ is the parallel translate to x of the sum of the contracted generalized eigenspaces of A, and the unstable bundle E^u at x is the translated sum of expanded eigenspaces.

In general, the invariant subbundles E^u and E^s of an Anosov diffeomorphism are integrable and tangent to a transverse pair of foliations $\mathcal{W}^u$ and $\mathcal{W}^s$, respectively (see, e.g. [12] for a proof of this). The leaves of $\mathcal{W}^s$ are uniformly contracted by f, and the leaves of $\mathcal{W}^u$ are uniformly contracted by f^{-1}. The leaves of these foliations are as smooth as f, but the tangent bundles to the leaves do not vary smoothly in the manifold. The regularity properties of these foliations play an important role in the ergodic properties of Anosov diffeomorphisms.

The first Anosov flows to be studied extensively were the geodesic flows for manifolds of negative sectional curvatures. As these flows are Hamiltonian, they are conservative. Eberhard Hopf showed in the 1930's that such geodesic flows for surfaces are ergodic with respect to Liouville measure [14]; it was not until the 1960's that ergodicity of all such flows was proved by Anosov [15]. The next section describes, in the context of Anosov diffeomorphisms, Hopf's method and important refinements due to Anosov and Sinai.

DA attractors. A simple way to produce a non-Anosov hyperbolic attractor on the torus is to start with an Anosov diffeomorphism, such as a linear hyperbolic automorphism, and deform it in a neighborhood of a fixed point, turning a saddle fixed point into a source, while preserving the stable foliation. If this procedure is carried out carefully enough, the resulting diffeomorphism is a dissipative hyperbolic diffeomorphism, called a *derived from Anosov (DA)* attractor. Other examples of hyperbolic attractors are the Plykin attractor and the solenoid. See [10].

Distortion estimates

Before describing the ergodic properties of hyperbolic systems, it is useful to pause for a brief discussion of distortion estimates. Distortion estimates are behind almost every result in smooth ergodic theory. In the hyperbolic setting, distortion estimates are applied to the action of f on unstable manifolds to show that the volume distortion of f along unstable manifolds can be controlled for arbitrarily many iterates.

The example mentioned at the end of the previous section illustrates the ideas in a distortion estimate. Suppose that $f\colon S^1 \to S^1$ is a C^2 expanding map, such as a C^2 small perturbation of T_2. Then there exist constants $\mu > 1$ and $C > 0$ such that $(f^n)'(x) > C\mu^n$ for all x and n.

Let d be the degree of f. If I is a sufficiently small open interval in S^1, then for each n, $f^{-n}(I)$ is a union of d disjoint intervals. Furthermore, each of these intervals has diameter at most $C^{-1}\mu^{-n}$ times the diameter of I. It is now possible to justify the assertion in Fundamental Principle #3 in this context.

Lemma 6.1. *There exists a constant $K \geq 1$ such that, for all $n \in \mathbb{N}$, and for all $x, y \in f^{-n}(I)$, one has:*

$$K^{-1} \leq \frac{(f^n)'(x)}{(f^n)'(y)} \leq K.$$

Proof. Since f is C^2 and f' is bounded away from 0, the function $\alpha(x) = \log(f'(x))$ is C^1. In particular, α is Lipschitz continuous: there exists a constant $L > 0$ such that, for all $x, y \in S^1$, $|\alpha(x) - \alpha(y)| < Ld(x, y)$. For $n \geq 0$, let $\alpha_n(x) = \log((f^n)'(x))$. The Chain Rule implies that $\alpha_n(x) = \sum_{i=0}^{n-1} \alpha(f^i(x))$.

The expanding hypothesis on f implies that for all $x, y \in f^n(I)$ and for $i = 0, \ldots, n$, one has $d(f^i(x), f^i(y)) \leq C^{-1}\mu^{n-i} d(f^n(x), f^n(y)) \leq C^{-1}\mu^{n-i}$. Hence

$$\begin{aligned}
|\alpha_n(x) - \alpha_n(y)| &\leq \sum_{i=0}^{n-1} |\alpha(f^i(x)) - \alpha(f^i(y))| \\
&\leq L \sum_{i=0}^{n-1} d(f^i(x), f^i(y)) \\
&\leq L \sum_{i=0}^{n-1} C^{-1}\mu^{n-i} \\
&< LC^{-1}\mu^{-1}(1 - \mu^{-1})^{-1}.
\end{aligned}$$

Setting $K = \exp(LC^{-1}\mu^{-1}(1 - \mu^{-1})^{-1})$, one now sees that $(f^n)'(x)/(f^n)'(y)$ lies in the interval $[K^{-1}.K]$, proving the claim. □

In this distortion estimate, the function $\alpha\colon M \to \mathbb{R}$ is called a *cocycle.* The same argument applies to any Lipschitz continuous (or even Hölder continuous) cocycle.

Ergodicity of expanding maps

The ergodic properties of C^2 expanding maps are completely understood. In particular, every conservative expanding map is ergodic, and every expanding map is conservative. The proofs of these facts use Fundamental Principles #1 and 3 in a fairly direct way.

Every C^2 conservative expanding map is ergodic with respect to volume. The proof is a straightforward adaptation of the proof of Proposition 3 (see, e.g. [2]). Here is a description of the proof for $M = S^1$. As remarked earlier, the proof of Proposition 3 adapts easily to a general expanding map $f\colon S^1 \to S^1$ once one shows that for every f-invariant set A, and every connected component J of $f^{-n}(S^1 \setminus \{p\})$, the quantity

$$\frac{\lambda(f^n(A \cap J) \colon f^n(J))}{\lambda(A \colon J)}$$

is bounded independently of n. This is a fairly direct consequence of the distortion estimate in Lemma 6.1 and is left as an exercise.

The same distortion estimates show that every C^2 expanding map is conservative, preserving a probability measure ν in the measure class of volume. Here is a sketch of the proof for the case $M = S^1$. To prove this, consider the push-forward $\lambda_n = f_*^n \lambda$. Then λ_n is equivalent to Lebesgue, and its Radon-Nikodym derivative $\frac{d\lambda_n}{d\lambda}$ is the density function

$$\rho_n(x) = \sum_{y \in f^{-n}(x)} \frac{1}{\mathrm{jac}_y f^n}.$$

Since $f_*^n\lambda$ is a probability measure, it follows that $\int_{S^1} \rho_n \, d\lambda = 1$. A simple argument using the distortion estimate above (and summing up over all d^n branches of f^{-n} at x) shows that there exists a constant $c \geq 1$ such that for all $x, y \in S^1$,

$$c^{-1} \leq \frac{\rho_n(x)}{\rho_n(y)} \leq c.$$

Since the integral of ρ_n is 1, the functions ρ_n are uniformly bounded away from 0 and ∞. It is easy to see that the measure $\nu_n = \frac{1}{n}\sum_{i=1}^n f_*^i \lambda$ has density $\frac{1}{n}\sum_{i=1}^n \rho_i$. Let ν be any subsequential weak* limit of ν_n; then ν is absolutely continuous, with density ρ bounded away from 0 and ∞. With a little more care, one can show that ρ is actually Lipschitz continuous.

As a passing comment, the ergodicity of ν and positivity of ρ imply that ν is the unique f-invariant measure absolutely continuous with respect to λ. With more work, one can show that ν is exact. See [2] for details.

Ergodicity of conservative Anosov diffeomorphisms

Like conservative C^2 expanding maps, conservative C^2 Anosov diffeomorphisms are ergodic. This subsection outlines a proof of this fact. Unlike expanding maps, however, Anosov diffeomorphisms need not be conservative. The subsection following this one describe a type of invariant measure that is "natural" with respect to volume, called a Sinai-Ruelle-Bowen (or SRB) measure. The central result for hyperbolic systems states that every hyperbolic attractor carries an SRB measure.

The Hopf Argument

In the 1930's Hopf [14] proved that the geodesic flow for a compact, negatively-curved surface is ergodic. His method was to study the Birkhoff averages of continuous functions along leaves of the stable and unstable foliations of the flow. This type of argument has been used since then in increasingly general contexts, and has come to be known as the Hopf Argument.

The core of the Hopf Argument is very simple. To any $f: M \to M$ one can associate the *stable equivalence relation* $\sim_s$, where $x \sim_s y$ iff $\lim_{n\to\infty} d(f^n(x), f^n(y)) = 0$. Denote by $W^s(x)$ the stable equivalence class containing x. When f is invertible, one defines the *unstable equivalence relation* to be the stable equivalence relation for f^{-1}, and one denotes by $W^u(x)$ the unstable equivalence class containing x.

The first step in the Hopf Argument is to show that Birkhoff averages for continuous functions are constant along stable and unstable equivalence classes. Let $\phi: M \to \mathbb{R}$ be an integrable function, and let

$$\overline{\phi} = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \phi \circ f^i. \tag{1}$$

Observe that if ϕ is continuous, then for every $x \in M$ and $x' \in W^s(x)$, $\lim_{n\to\infty} |\phi(f^i(x)) - \phi(f^i(x'))| = 0$. It follows immediately that $\overline{\phi}_f(x) = \overline{\phi}_f(x')$. In particular, if the limit in (1) exists at x, then it exists and is constant on $W^s(x)$.

Fundamental Principle #4: *Birkhoff averages of continuous functions are constant along stable equivalence classes.*

The next step of Hopf's argument confines itself to the situation where f is conservative and Anosov. In this case, f is invertible, the stable equivalence classes are precisely the leaves of the stable foliation $\mathcal{W}^s$, and the unstable equivalence classes are the leaves of the unstable foliation $\mathcal{W}^u$. Since f is conservative, the Ergodic Theorem implies that for every L^2 function ϕ, the function $\overline{\phi}_f$ is equal (mod 0) to the projection of ϕ onto the f-invariant functions in L^2. Since this projection is continuous, and the continuous functions are dense in L^2, to prove that f is ergodic, it suffices to show that the projection of any continuous function is trivial. That is, it suffices to show that for every continuous ϕ, the function $\overline{\phi}_f$ is constant (a.e.).

To this end, let $\phi: M \to \mathbb{R}$ be continuous. Since the f-invariant functions coincide with the f^{-1}-invariant functions, one obtains that $\overline{\phi}_f = \overline{\phi}_{f^{-1}}$ a.e. The previous argument shows $\overline{\phi}_f$ is constant along $\mathcal{W}^s$-leaves and $\overline{\phi}_{f^{-1}}$ is constant along $\mathcal{W}^u$-leaves. The desired conclusion is that $\overline{\phi}_f$ is a.e. constant. It suffices to show this in a local chart, since the manifold M is connected. In a local chart, after a smooth change of coordinates, one obtains a pair of transverse foliations $\mathcal{F}_1$, $\mathcal{F}_2$ of the cube $[-1,1]^n$ by disks, and a measurable function $\psi: [-1,1]^n \to \mathbb{R}$ that is constant along the leaves of $\mathcal{F}_1$ and constant along the leaves of $\mathcal{F}_2$.

When the foliations $\mathcal{F}_1$ and $\mathcal{F}_2$ are smooth (at least C^1), one can perform a further smooth change of coordinates so that $\mathcal{F}_1$ and $\mathcal{F}_2$ are transverse coordinate subspace foliations. In this case, Fubini's theorem implies that any measurable function that is constant along two transverse coordinate foliations is a.e. constant. This completes the proof in the case that the foliations $\mathcal{W}^s$ and $\mathcal{W}^u$ are smooth. In Hopf's original argument, the stable and unstable foliations were assumed to be C^1 foliations (a hypotheses satisfied in the examples he considered, due to low-dimensionality. See also [16], where a pinching condition on the curvature, rather than low dimensionality, implies this C^1 condition on the foliations.)

Absolute continuity

For a general Anosov diffeomorphism or flow, the stable and unstable foliations are not C^1, and so the final step in Hopf's orginal argument does not apply. The fundamental advance of Anosov and Anosov-Sinai was to prove that the stable and unstable foliations of an Anosov diffeomorphism (conservative or not) satisfy a weaker condition than smoothness, called *absolute continuity*. For conservative systems, absolute continuity is enough to finish Hopf's argument, proving that every C^2 conservative Anosov diffeomorphism is ergodic [15, 17].

For a definition and careful discussion of absolute continuity of a foliation $\mathcal{F}$, see [13]. Two consequences of the absolute continuity of $\mathcal{F}$ are:

(AC1) If $A \subset M$ is any measurable set, then

$$\lambda(A) = 0 \quad \Longleftrightarrow \quad \lambda_{\mathcal{F}(x)}(A) = 0, \quad \text{for} \quad \lambda - \text{a.e. } x \in M,$$

where $\lambda_{\mathcal{F}(x)}$ denotes the induced Riemannian volume on the leaf of $\mathcal{F}$ through x.

(AC2) If τ is any small, smooth disk transverse to a local leaf of $\mathcal{F}$, and $T \subset \tau$ is a 0-set in τ (with respect to the induced Riemannian volume on τ), then the union of the $\mathcal{F}$ leaves through points in T has Lebesgue measure 0 in M.

The proof that $\mathcal{W}^s$ and $\mathcal{W}^u$ are absolutely continuous has a similar flavor to the proof that an expanding map has a unique absolutely continuous invariant measure (although the cocycles involved are Hölder continuous, rather than Lipschitz), and the facts are intimately related.

With absolute continuity of the stable and unstable foliations in hand, one can now prove:

Theorem 6.2. *(Anosov) Let f be a C^2, conservative Anosov diffeomorphism. Then f is ergodic.*

Sketch of proof. By the Hopf Argument, it suffices to show that if ψ^s and ψ^u are L^2 functions with the following properties:

(1) ψ^s is constant along leaves of $\mathcal{W}^s$,
(2) ψ^u is constant along leaves of $\mathcal{W}^u$, and
(3) $\psi^s = \psi^u$ a.e.,

then ψ^s (and so ψ^u as well) is constant a.e.

This is proved using the absolute continuity of $\mathcal{W}^u$ and $\mathcal{W}^s$. Since M is connected, one may argue this locally. Let G be the full measure set of $p \in M$ such that $\psi^s = \psi^u$. Absolute continuity of $\mathcal{W}^s$ (more precisely, consequence (AC1) of absolute continuity described above) implies that for almost every $p \in M$, G has full measure in $\mathcal{W}^s(p)$. Pick such a p. Then for almost every $q \in \mathcal{W}^s(p)$, $\psi^s(q) = \psi^u(p)$; defining G' to be the union over all $q \in \mathcal{W}^s(p) \cap G$ of $\mathcal{W}^u(q)$, one obtains that ψ^s is constant on $G \cap G'$. But now, since $\mathcal{W}^s(p) \cap G$ has full measure in $\mathcal{W}^s(p)$, the absolute continuity of $\mathcal{W}^u$ (consequence (AC2) above) implies that G' has full measure in a neighborhood of p. Hence ψ^s is a.e. constant in a neighborhood of p, completing the proof. $\square$

SRB measures

In the absence of a smooth invariant measure, it is still possible for a map to have an invariant measure that behaves naturally with respect to volume. In computer simulations one observes such measures when one picks a point x at random and plots many iterates of x; in many systems, the resulting picture is surprisingly insensitive to the initial choice of x. What appears to be happening in these systems is that the trajectory of almost every x in an open set U is converging to the support of a singular invariant probability measure μ. Furthermore, for any open set V, the proportion of forward iterates of x spent in V appears to converge to $\mu(V)$ as the number of iterates tends to ∞.

In the 1960's and 70's, Sinai, Ruelle and Bowen rigorously established the existence of these physically observable measures for hyperbolic attractors [18, 19, 20]. Such measures are now known as Sinai-Ruelle-Bowen (SRB) measures, and have been shown to exist for non-hyperbolic maps with some hyperbolic features. Yhis subsection describes the construction of SRB measures for hyperbolic attractors.

An f-invariant probability measure μ is called an *SRB (or physical) measure* if there exists an open set $U \subset M$ containing the support of μ such that, for every continuous function $\phi\colon M \to \mathbb{R}$ and λ-a.e. $x \in U$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \phi(f^i(x)) = \int_M f\, d\mu.$$

The maximal open set U with this property is called the *basin* of f. To exclude the possibility that the SRB measure is supported on a periodic sink, one often adds the condition that at least one of the Lyapunov exponents of f with respect μ is positive. Other definitions of SRB measure have been proposed (see [21]). Note that every ergodic absolutely continuous invariant measure with positive density in an open set is an SRB measure. Note also that an SRB measure for f is not in general an SRB measure for f^{-1}, unless f preserves an ergodic absolutely continuous invariant measure.

Every transitive Anosov diffeomorphism carries a unique SRB measure. To prove this, one defines a sequence of probability measures ν_n on M as follows. Fix a point $p \in M$, and define ν_0 to be the normalized restriction of Riemannian volume to a ball B^u in $\mathcal{W}^u(p)$. Set $\nu_n = \frac{1}{n}\sum_{i=1}^{n} f^i_*\nu_0$. Distortion estimates show that the density of ν_n on its support inside $\mathcal{W}^u$ is bounded, above and below, independently of n. Passing to a subsequential weak* limit, one obtains a probability measure ν on M with bounded densities on $\mathcal{W}^u$-leaves.

To show that ν is an SRB measure, choose a point $q \in M$ in the support of ν. Since ν has positive density on unstable manifolds, almost every point in a neighborhood of q in $\mathcal{W}^u(q)$ is a regular point for f (that is, a point where the forward Birkhoff averages of every continuous function exist). A variation on the Hopf Argument, using the absolute continuity of $\mathcal{W}^s$, shows that ν is an ergodic SRB measure.

A similar argument shows that every transitive hyperbolic attractor admits an ergodic SRB measure. In fact this SRB measure has much stronger mixing properties, namely, it is Bernoulli. To prove this, one first constructs a Markov partition [22] conjugating the action of f to a Bernoulli shift. This map sends the SRB measure to a Gibbs state for a mixing Markov shift (see the entry on Equilibrium States in Ergodic Theory in this volume). A result that subsumes all of the results in this section is:

Theorem 6.3. *(Sinai, Ruelle, Bowen) Let $\Lambda \subset M$ be a transitive hyperbolic attractor for a C^2 map $f\colon M \to M$. Then f has an ergodic SRB measure μ supported on Λ. Moreover: the disintegration of μ along unstable manifolds of Λ is equivalent* *to the induced Riemannian volume, the Lyapunov exponents of μ are all positive, and μ is Bernoulli.*

7. Beyond uniform hyperbolicity

The methods developed in the smooth ergodic theory of hyperbolic maps have been extended beyond the hyperbolic context. Two natural generalizations of hyperbolicity are:

- partial hyperbolicity, which requires uniform expansion of E^u and uniform contraction of E^s, but allows central directions at each point, in which the expansion and contraction is dominated by the behavior in the hyperbolic directions; and
- nonuniform hyperbolicity, which requires hyperbolicity along almost every orbit, but allows the expansion of E^u and the contraction of E^s to weaken near the exceptional set where there is no hyperbolicity.

This section discusses both generalizations.

Partial hyperbolicity

Brin and Pesin [23] and independently Pugh and Shub [24] first examined the ergodic properties of partially hyperbolic systems soon after the work of Anosov and Sinai on hyperbolic systems. One says that a diffeomorphism $f\colon M \to M$ of a compact manifold M is *partially hyperbolic* if there is a nontrivial, continuous splitting of the tangent bundle, $TM = E^s \oplus E^c \oplus E^u$, invariant under Df, such that E^s is uniformly contracted, E^u is uniformly expanded, and E^c is dominated, meaning that for some $n \geq 1$ and for all $x \in M$:

$$\|D_x f^n|_{E^s}\| < m(D_x f^n|_{E^c}) \leq \|D_x f^n|_{E^c}\| < m(D_x f^n|_{E^u}).$$

Partial hyperbolicity is a C^1-open condition: any diffeomorphism sufficiently C^1-close to a partially hyperbolic diffeomorphism is itself partially hyperbolic. For an extensive discussion of examples of partially hyperbolic dynamical systems, see the survey article [25] and the book [26]. Among these examples are: the time-1 map of an Anosov flow, the frame flow for a compact manifold of negative sectional curvature, and many affine transformations of compact homogeneous spaces. All of these examples preserve the volume induced by a Riemannian metric on M.

As in the Anosov case, the stable and unstable bundles E^s and E^u of a partially hyperbolic diffeomorphism are tangent to foliations, again denoted by $\mathcal{W}^s$ and $\mathcal{W}^u$ respectively [23]. Brin-Pesin and Pugh-Shub proved that these foliations are absolutely continuous.

A partially hyperbolic diffeomorphism $f\colon M \to M$ is *accessible* if any point in M can be reached from any other along an *su-path*, which is a concatenation of finitely many subpaths, each of which lies entirely in a single leaf of $\mathcal{W}^s$ or a single leaf of $\mathcal{W}^u$. Accessibility is a global, topological property of the foliations $\mathcal{W}^u$ and $\mathcal{W}^s$ that is the analogue of transversality of $\mathcal{W}^u$ and $\mathcal{W}^s$ for Anosov diffeomorphisms. In fact, the transversality of these foliations in the Anosov case immediately implies that every Anosov diffeomorphism is accessible. Fundamental Principle #4 suggests that accessibility might be related to ergodicity for conservative systems.

Conservative partially hyperbolic diffeomorphisms

Motivated by a breakthrough result with Grayson [27], Pugh and Shub conjectured that accessibility implies ergodicity, for a C^2, partially hyperbolic conservative diffeomorphism [28]. This conjecture has been proved under the hypothesis of center bunching [29], which is a mild spectral condition on the restriction of Df to the center bundle E^c. Center bunching is satisfied by most examples of interest, including all partially hyperbolic diffeomorphisms with $\dim(E^c) = 1$. The proof in [29] is a modification of the Hopf Argument using Lebesgue density points and a delicate analysis of the geometric and measure-theoretic properties of the stable and unstable foliations.

In the same article, Pugh and Shub also conjectured that accessibility is a widespread phenomenon, holding for an open and dense set (in the C^r topology) of partially hyperbolic diffeomorphisms. This conjecture has been proved completely for $r = 1$ [30], and for all r, with the additional assumption that the central bundle E^c is one dimensional [31].

Together, these two conjectures imply the third, central conjecture: in [28]:

Conjecture 7.1 (Pugh–Shub). For any $r \geq 2$, the C^r, conservative partially hyperbolic diffeomorphisms contain a C^r open and dense set of ergodic diffeomorphisms.

The validity of this conjecture in the absence of center bunching is currently an open question.

Dissipative partially hyperbolic diffeomorphisms

There has been some progress in constructing SRB-type measures for dissipative partially hyperbolic diffeomorphisms, but the theory is less developed than in the conservative case. Using the same construction as for Anosov diffeomorphisms, one can construct invariant probability measures that are smooth along the $\mathcal{W}^u$ foliation [32]. Such measures are referred to as u-Gibbs measures. Since the stable bundle E^s is not transverse to the unstable bundle E^u, the Anosov argument cannot be carried through to show that u-Gibbs measures are SRB measures.

Nonetheless, there are conditions that imply that a u-Gibbs measure is an SRB measure: for example, a u-Gibbs measure is SRB if it is the unique u-Gibbs measure [33], if the bundle $E^s \oplus E^c$ is nonuniformly contracted [34], or if the bundle $E^u \oplus E^c$ is nonuniformly expanded [35]. The proofs of the latter two results use Pesin Theory, which is explained in the next subsection.

SRB measures have also been constructed in systems where E^c is nonuniformly hyperbolic [36], and in (noninvertible) partially hyperbolic covering maps where E^c is 1-dimensional [37]. It is not known whether accessibility plays a role in the existence of SRB measures for dissipative, non-Anosov partially hyperbolic diffeomorphisms.

Nonuniform hyperbolicity

The concept of Lyapunov exponents gives a natural way to extend the notion of hyperbolicity to systems that behave hyperbolically, but in a nonuniform manner. The fundamental principles described above, suitably modified, apply to these nonuniformly hyperbolic systems and allow for the development of a smooth ergodic theory for these systems. This program was initially proposed and carried out by Yakov Pesin in the 1970's [38] and has come to be known as *Pesin theory.*

Oseledec's Theorem implies that if a smooth map f satisfying the condition $m(D_xf) > 0$ preserves a probability measure ν, then for ν-a.e. $x \in M$ and every nonzero vector $v \in T_xM$, the limit

$$\lambda(x, v) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log \|D_xf^i(v)\|$$

exists. The number $\lambda(x, v)$ is called the *Lyapunov exponent at x in the direction of v*. For each such x, there are finitely many possible values for the exponent $\lambda(x, v)$, and the function $x \mapsto \lambda(x, \cdot)$ is measurable. See the discussion of Oseledec's Theorem in the entry on Ergodic Theorems in this volume.

Let f be a smooth map. An f-invariant probability measure μ is *hyperbolic* if the Lyapunov exponents of μ-a.e. point are all nonzero. Observe that any invariant measure of a hyperbolic map is a hyperbolic measure.

A conservative diffeomorphism $f\colon M \to M$ is *nonuniformly hyperbolic* if the invariant measure equivalent to volume is hyperbolic. The term "nonuniform" is a bit misleading, as uniformly hyperbolic conservative systems are also nonuniformly hyperbolic. Unlike uniform hyperbolicity, however, nonuniform hyperbolicity allows for the *possibility* of different strengths of hyperbolicity along different orbits.

Nonuniformly hyperbolic diffeomorphisms exist on all manifolds [39, 40], and there are C^1- open sets of nonuniformly hyperbolic diffeomorphisms that are not Anosov diffeomorphisms [41]. In general, it is a very difficult problem to establish whether a given map carries a hyperbolic measure that is nonsingular with respect to volume.

Hyperbolic blocks

As mentioned above, the derivative of f along almost every orbit of a nonuniformly hyperbolic system looks like the derivative down the orbit of a uniformly hyperbolic system; the nonuniformity can be detected only by examining a positive measure set of orbits. Recall that Lusin's Theorem in measure theory states that every Borel measurable function is continuous on the complement of an arbitrarily small measure set. A sort of analogue of Lusin's theorem holds for nonuniformly hyperbolic maps: every C^2, nonuniformly hyperbolic diffeomorphism is uniformly hyperbolic on a (noninvariant) compact set whose complement has arbitrarily small measure. The precise formulation of this statement is omitted, but here are some of its salient features.

If μ is a hyperbolic measure for a C^2 diffeomorphism, then attached to μ-a.e. point $x \in M$ are transverse, smooth stable and unstable manifolds for f. The collection of all stable manifolds is called the *stable lamination* for f, and the collection of all unstable manifolds is called the *unstable lamination* for f. The stable lamination is invariant under f, meaning that f sends the stable manifold at x into the stable manifold for $f(x)$. The stable manifold through x is contracted uniformly by all positive iterates of f in a neighborhood of x. Analogous statements hold for the unstable manifold of x, with f replaced by f^{-1}.

The following quantites vary measurably in $x \in M$:

- the (inner) radii of the stable and unstable manifolds through x,
- the angle between stable and unstable manifolds at x, and
- the rates of contraction in these manifolds.

The stable and unstable laminations of a nonuniformly hyperbolic system are absolutely continuous. The precise definition of absolute continuity here is slightly different than in the uniformly and partially hyperbolic setting, but the consequences (AC1) and (AC2) of absolute continuity continue to hold.

Ergodic properties of nonuniformly hyperbolic diffeomorphisms

Since the stable and unstable laminations are absolutely continuous, the Hopf Argument can be applied in this setting to show:

Theorem 7.2 (Pesin)**.** *Let f be C^2, conservative and nonuniformly hyperbolic. Then there exists a (mod 0) partition $\mathcal{P}$ of M into countably many f-invariant sets of positive volume such that the restriction of f to each $P \in \mathcal{P}$ is ergodic.*

The proof of this theorem is also exposited in [42]. The countable partition can in examples be countably infinite; nonuniform hyperbolicity alone does not imply ergodicity.

The dissipative case

As mentioned above, establishing the existence of a nonsingular hyperbolic measure is a difficult problem in general. In systems with some global form of hyperbolicity, such as partial hyperbolicity, it is sometimes possible to "borrow" the expansion from the unstable direction and lend it to the central direction, via a small perturbation. Nonuniformly hyperbolic attractors have been constructed in this way [43]. This method is also behind the construction of a C^1 open set of nonuniformly hyperbolic diffeomorphisms in [41].

For a given system of interest, it is sometimes possible to prove that a given invariant measure is hyperbolic by establishing an *approximate* form of hyperbolicity. The idea, due to Wojtkowski and called the *cone method*, is to isolate a measurable bundle of cones in TM defined over the support of the measure, such that the cone at a point x is mapped by D_xf into the cone at $f(x)$. Intersecting the images of these cones under all iterates of Df, one obtains an invariant subbundle of TM over the support of f that is nonuniformly expanded.

Lai-Sang Young has developed a very general method [44] for proving the existence of SRB measures with strong mixing properties in systems that display "some hyperbolicity." The idea is to isolate a region X in the manifold where the first return map is hyperbolic and distortion estimates hold. If this can be done, then the map carries a mixing, hyperbolic SRB measure. The precise rate of mixing is determined by the properties of the return-time function to X; the longer the return times, the slower the rate of mixing.

More results on the existence of hyperbolic measures are discussed in the next section.

An important subject in smooth ergodic theory is the relationship between entropy, Lyapunov exponents, and dimension of invariant measures of a smooth map. Significant results in this area include the Pesin entropy formula [45], the Ruelle entropy inequality, [46], the entropy-exponents-dimension formula of Ledrappier-Young [47, 48], and the proof by Barreira-Pesin-Schmeling that hyperbolic measures have a well-defined dimension [49]. The entry on Hyperbolic Dynamical Systems in this volume contains a discussion of these results; see this entry there for further information.

8. The presence of critical points and other singularities

Now for a discussion of the aforementioned technical difficulties that arise in the presence of singularities and critical points for the derivative.

Singularities, that is, points where Df (or even f) fails to be defined, arise naturally in the study of billiards and hard sphere gases. The first subsection discusses some progress made in smooth ergodic theory in the presence of singularities.

Critical points, that is, points where Df fails to be invertible, appear inescapably in the study of noninvertible maps. This type of complication already shows up for noninvertible maps in dimension 1, in the study of unimodal maps of the interval. The second subsection discusses the technique of parameter exclusion, developed by Jakobson, which allows for an ergodic analysis of a parametrized family of maps with criticalities.

The technical advances used to overcome these issues in the interval have turned out to have applications to dissipative, nonhyperbolic, diffeomorphisms in higher dimension, where the derivative is "nearly critical" in places. The last subsection describes extensions of the parameter exclusion technique to these near-critical maps.

Hyperbolic billiards and hard sphere gases

In the 1870's the physicist Ludwig Boltzmann hypothesized that in a mechanical system with many interacting particles, physical measurements (observables), averaged over time, will converge to their expected value as time approaches infinity. The underlying dynamical system in this statement is a Hamiltonian system with many degrees of freedom, and the "expected value" is with respect to Liouville measure. Loosely phrased in modern terms, Boltzmann's hypothesis states that a generic Hamiltonian system of this form will be ergodic on constant energy submanifolds. Reasoning that the time scales involved in measurement of an observable in such a system are much larger than the rate of evolution of the system, Boltzmann's hypothesis allowed him to assume that physical quantities associated to such a system behave like constants.

In 1963, Sinai revived and formalized this ergodic hypothesis, stating it in a concrete formulation known as the Boltzmann-Sinai Ergodic Hypothesis. In Sinai's formulation, the particles were replaced by N hard, elastic spheres, and to compactify the problem, he situated the spheres on a k-torus, $k = 2, 3$. The Boltzmann-Sinai Ergodic Hypothesis is the conjecture that the induced Hamiltonian system on the $2kN$-dimensional configuration space is ergodic on constant energy manifolds, for any $N \geq 2$.

Sinai verified this conjecture for $N = 2$ by reducing the problem to a billiard map in the plane. As background for Sinai's result, a brief discussion of planar billiard maps follows.

Let $D \subset \mathbb{R}^s$ be a connected region whose boundary ∂D is a collection of closed, piecewise smooth simple curves the plane. The *billiard map* is a map defined (almost everywhere) on $\partial D \times [-\pi, \pi]$. To define this map, one identifies each point $(x, \theta) \in \partial D \times [-\pi, \pi]$ with an inward-pointing tangent vector at x in the plane, so that the normal vector to ∂D at x corresponds to the pair $(x, \pi/2)$. This can be done in a unique way on the smooth components of ∂D. Then $f(x, \theta)$ is obtained by following the ray originating at (x, θ) until it strikes the boundary ∂D for the first time at (x', θ'). Reflecting this vector about the normal at x', define $f(x, \theta) = (x', \pi - \theta')$.

It is not hard to see that the billiard map is conservative. The billiard map is piecewise smooth, but not in general smooth: the degree of smoothness of f is one less than the degree of smoothness of ∂D. In addition to singularities arising from the corners of the table, there are singularities arising in the *second* derivative of f at the tangent vectors to the boundary.

In the billiards studied studied by Sinai, the boundary ∂D consists of a union of concave circular arcs and straight line segments. Similar billiards, but with convex circular arcs, were first studied by Bunimovich [51]. Sinai and Bunimovich proved that these billiards are ergodic and nonuniformly hyperbolic. For the Boltzmann-Sinai problem with $N \geq 3$, the relevant associated dynamical system is a higher dimensional billiard table in euclidean space, with circular arcs replaced by cylindrical boundary components.

In a planar billiard table with circular/flat boundary, the behavior of vectors encountering a flat segment of boundary is easily understood, as is the behavior of vectors meeting a circular segment in a neighborhood of the normal vector. If the billiard map is ergodic, however, every open set of vectors will meet the singularities in the table infinitely many times. To establish the nonuniform hyperbolicity of such billiard tables via conefieds, it is therefore necessary to understand precisely the fraction of time orbits spend near these singularities. Furthermore, to use the Hopf argument to establish ergodicity, one must avoid the singularities in the second derivative, where distortion estimates break down. The techniques for overcoming these obstacles involve imposing restrictions on the geometry of the table (even more so for higher dimensional tables), and are well beyond the scope of this paper.

The study of hyperbolic billiards and hard sphere gases has a long and involved history. See the articles [50] and [52] for a survey of some of the results and techniques in the area. A discussion of methods in singular smooth ergodic theory, with particular applications to the Lorentz attractor, can be found in [53]. Another, more classical, reference is [54], which contains a formulation of properties on a critical set, due to Katok-Strelcyn, that are useful in establishing ergodicity of systems with singularities.

Interval maps and parameter exclusion

The logistic family of maps $f_t \colon x \mapsto tx(1-x)$ defined on the interval $[0,1]$ is very simple to define but exhibits an astonishing variety of dynamical features as the parameter t varies. For small positive values of t, almost every point in I is attracted under the map f_t to the sink at -1. For values of $t > 4$, the map has a repelling hyperbolic Cantor set. As the value of t increases between 0 and 4, the map f_t undergoes a cascade of period-doubling bifurcations, in which a periodic sink of period 2^n becomes repelling and a new sink of period 2^{n+1} is born. At the accumulation of period doubling at $t \approx 3.57$, a periodic point of period 3 appears, forcing the existence of periodic points of all periods. The dynamics of f_t for t close to 4 has been the subject of intense inquiry in the last 20 years.

The map f_t, for t close to 4, shares some of the features of the doubling map T_2; it is 2-to-1, except at the critical point $\frac{1}{2}$, and it is uniformly expanding in the complement of a neighborhood of this critical point. Because this neighborhood of the critical point is not invariant, however, the only invariant sets on which f_t is uniformly hyperbolic have measure zero. Furthermore, the second derivative of f_t

vanishes at the critical point, making it impossible to control distortion for orbits that spend too much time near the critical point.

Despite these serious obstacles, Michael Jakobson [55] found a method for constructing absolutely continuous invariant measures for maps in the logistic family. The method has come to be known as *parameter exclusion* and has seen application far beyond the logistic family. As with billiards, it is possible to formulate geometric conditions on the map f_t that control both expansion (hyperbolicity) and distortion on a positive measure set. As these conditions involve understanding infinitely many iterates of f_t, they are impossible to verify for a given parameter value t.

Using an inductive formulation of this condition, Jakobson showed that the set of parameters t near 4 that fail to satisfy the condition at iterate n have exponentially small measure (in n). He thereby showed that for a positive Lebesgue measure set of parameter values t, the map f_t has an absolutely continuous invariant measure [55]. This measure is ergodic (mixing) and has a positive Lyapunov exponent. The delicacy of Jakobson's approach is confirmed by the fact that for an open and dense set of parameter values, almost every orbit is attracted to a periodic sink, and so f_t has no absolutely continuous invariant measure [56, 57]. Jakobson's method applies not only to the logistic family but to a very general class of C^3 one-parameter families of maps on the interval.

Near-critical diffeomorphisms

Jakobson's method in one dimension proved to extend to certain highly dissipative diffeomorphisms. The seminal paper in this extension is due to Benedicks and Carleson; the method has since been extended in a series of papers [59, 60, 61] and has been formulated in an abstract setting [62].

This extension turns out to be highly nontrivial, but it is possible to describe informally the similarities between the logistic family and higher-dimensional "near critical" diffeomorphisms. The diffeomorphisms to which this method applies are crudely hyperbolic with a one dimensional unstable direction. Roughly this means that in some invariant region of the manifold, the image of a small ball under f will be stretched significantly in one direction and shrunk in all other directions. The directions of stretching and contraction are transverse in a large proportion of the invariant region, but there are isolated "near critical" subregions where expanding and contracting directions are nearly tangent.

The dynamics of such a diffeomorphism are very close to 1-dimensional if the contraction is strong enough, and the diffeomorphism resembles an interval map with isolated critcal points, the critical points corresponding to the critical regions where stable and unstable directions are tangent.

An illustration of this type of dynamics is the Hénon family of maps $f_{a,b}\colon (x,y) \mapsto (1-ax^2+by, x)$, the original object of study in Benedicks-Carlesson's work. When the parameter b is set to 0, the map $f_{a,b}$ is no longer a diffeomorphism, and indeed is precisely a projection composed with the logisitic map. For small values of b and appropriate values of a, the Hénon map is strongly dissipative and displays the near critical behavior described in the previous paragraph. In analogy to Jakobson's result, there is a positive measure set of parameters near $b = 0$ where $f_{a,b}$ has a mixing, hyperbolic SRB measure.

See [63] for a detailed exposition of the parameter exclusion method for Hénon-like maps.

9. Future directions

In addition to the open problems discussed in the previous sections, there are several general questions and problems worth mentioning:

- What can be said about systems with everywhere vanishing Lyapunov exponents? Open sets of such systems exist in arbitrary dimension. Pesin theory carries into the nonuniformly hyperbolic setting the basic principles from uniformly hyperbolic theory (in particular, Fundamental Principles #3 and 4 above). To what extent do properties of isometric and unipotent systems (for example, Fundamental Principle #2) extend to conservative systems all of whose Lyaponov exponents vanish?
- Can one establish the existence of and analyze in general the conservative systems on surfaces that have two positive measure regimes: one where Lyapunov exponents vanish, and the other where they are nonzero? Such systems are conjectured exist in the presence of KAM phenomena surrounding elliptic periodic points.
- On a related note, how common are conservative systems whose Lyapunov exponents are nonvanishing on a positive measure set? See [64] for a discussion.
- Find a broad description of those dissipative systems that admit finitely (or countably) many physical measures. Are such systems dense among all dissipative systems, or possibly generic among a restricted class of systems? See [41, 65] for several questions and conjectures related to this problem.
- Extend the methods in the study of systems with singularities to other specific systems of interest, including the infinite dimensional systems that arise in the study of partial differential equations.
- Carry the methods of smooth ergodic theory further into the study of smooth actions of discrete groups (other than the integers) on manifolds. When do such actions admit (possibly non-invariant) "physical" measures?

There are other interesting open areas of future inquiry, but this gives a good sample of the range of possibilities.

References

[1] Arnol'd, V. I.; Avez, A. *Ergodic problems of classical mechanics.* Translated from the French by A. Avez. W. A. Benjamin, Inc., New York-Amsterdam 1968
[2] Mañé, Ricardo *Ergodic theory and differentiable dynamics.* Translated from the Portuguese by Silvio Levy. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 8. Springer-Verlag, Berlin, 1987.
[3] Baladi, Viviane *Positive transfer operators and decay of correlations.* Advanced Series in Nonlinear Dynamics, 16. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.
[4] Kifer, Yuri *Ergodic theory of random transformations.* Progress in Probability and Statistics, 10. Birkhuser Boston, Inc., Boston, MA, 1986.
[5] Kifer, Yuri *Random perturbations of dynamical systems.* Progress in Probability and Statistics, 16. Birkhuser Boston, Inc., Boston, MA, 1988.
[6] Liu, Pei-Dong; Qian, Min *Smooth ergodic theory of random dynamical systems.* Lecture Notes in Mathematics, 1606. Springer-Verlag, Berlin, 1995.
[7] Bonatti, Christian; Dìaz, Lorenzo J.; Viana, Marcelo *Dynamics beyond uniform hyperbolicity. A global geometric and probabilistic perspective.* Encyclopaedia of Mathematical Sciences, 102. Mathematical Physics, III. Springer-Verlag, Berlin, 2005.

[8] Cornfeld, I. P.; Fomin, S. V.; Sinaĭ, Ya. G. *Ergodic theory.* Translated from the Russian by A. B. Sosinskiĭ. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 245. Springer-Verlag, New York, 1982.

[9] Hirsch, Morris W. *Differential topology.* Graduate Texts in Mathematics, No. 33. Springer-Verlag, New York-Heidelberg, 1976.

[10] Robinson, Clark *Dynamical systems.* Stability, symbolic dynamics, and chaos. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.

[11] Katok, Anatole; Hasselblatt, Boris *Introduction to the modern theory of dynamical systems.* With a supplementary chapter by Katok and Leonardo Mendoza. Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge, 1995.

[12] Hirsch, M. W.; Pugh, C. C.; Shub, M. *Invariant manifolds.* Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977.

[13] Brin, Michael; Stuck, Garrett *Introduction to dynamical systems.* Cambridge University Press, Cambridge, 2002.

[14] Hopf, Eberhard *Statistik der geodtischen Linien in Mannigfaltigkeiten negativer Krummung.* (German) Ber. Verh. Sachs. Akad. Wiss. Leipzig 91, (1939). 261–304.

[15] Anosov, D. V. *Geodesic flows on closed Riemann manifolds with negative curvature.* Proceedings of the Steklov Institute of Mathematics, No. 90 (1967). Translated from the Russian by S. Feder American Mathematical Society, Providence, R.I. 1969

[16] Sinaĭ, Ja. G. *Geodesic flows on compact surfaces of negative curvature.* Dokl. Akad. Nauk SSSR 136 549–552 (Russian); translated as Soviet Math. Dokl. 2 1961 106–109.

[17] Anosov, D. V.; Sinaĭ, Ja. G. *Certain smooth ergodic systems.* (Russian) Uspehi Mat. Nauk 22 1967 no. 5 (137), 107–172.

[18] Sinaĭ, Ja. G. *Gibbs measures in ergodic theory.* (Russian) Uspehi Mat. Nauk 27 (1972), no. 4(166), 21–64.

[19] Ruelle, David *A measure associated with axiom-A attractors.* Amer. J. Math. 98 (1976), no. 3, 619–654.

[20] Bowen, Rufus *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, 1975.

[21] Young, Lai-Sang *What are SRB measures, and which dynamical systems have them?* Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. J. Statist. Phys. 108 (2002), no. 5-6, 733–754.

[22] Bowen, Rufus *Markov partitions for Axiom* A *diffeomorphisms.* Amer. J. Math. 92 1970 725–747.

[23] Brin, M. I.; Pesin, Ja. B. *Partially hyperbolic dynamical systems.* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 38 (1974), 170–212.

[24] Pugh, Charles; Shub, Michael *Ergodicity of Anosov actions.* Invent. Math. 15 (1972), 1–23.

[25] Burns, Keith; Pugh, Charles; Shub, Michael; Wilkinson, Amie *Recent results about stable ergodicity.* Smooth ergodic theory and its applications (Seattle, WA, 1999), 327–366, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[26] Pesin, Yakov B. *Lectures on partial hyperbolicity and stable ergodicity.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zrich, 2004.

[27] Grayson, Matthew; Pugh, Charles; Shub, Michael *Stably ergodic diffeomorphisms.* Ann. of Math. (2) 140 (1994), no. 2, 295–329.

[28] Pugh, Charles; Shub, Michael *Stable ergodicity and partial hyperbolicity.* International Conference on Dynamical Systems (Montevideo, 1995), 182–187, Pitman Res. Notes Math. Ser., 362, Longman, Harlow, 1996.

[29] Burns, K; Wilkinson, A. *On the ergodicity of partially hyperbolic systems.* Ann. of Math., to appear.

[30] Dolgopyat, Dmitry; Wilkinson, Amie *Stable accessibility is C^1 dense.* Geometric methods in dynamics. II. Astrisque No. 287 (2003).

[31] Rodríguez Hertz, A., Rodríguez Hertz, F., and R. Ures, *Partially hyperbolic systems with 1D-center bundle: I.– stable ergodicity.* Invent. Math., to appear.

[32] Pesin, Ya. B.; Sinaĭ, Ya. G. *Gibbs measures for partially hyperbolic attractors.* Ergodic Theory Dynam. Systems 2 (1982), no. 3-4, 417–438 (1983).

[33] Dolgopyat, Dmitry *On differentiability of SRB states for partially hyperbolic systems.* Invent. Math. 155 (2004), no. 2, 389–449.

[34] Bonatti, Christian; Viana, Marcelo *SRB measures for partially hyperbolic systems whose central direction is mostly contracting.* Israel J. Math. 115 (2000), 157–193.

[35] Alves, José F.; Bonatti, Christian; Viana, Marcelo *SRB measures for partially hyperbolic systems whose central direction is mostly expanding.* Invent. Math. 140 (2000), no. 2, 351–398.

[36] Burns, Keith; Dolgopyat, Dmitry; Pesin, Yakov; Pollicott Mark *Stable ergodicity for partially hyperbolic attractors with negative central exponents.* preprint.

[37] Tsujii, Masato *Physical measures for partially hyperbolic surface endomorphisms.* Acta Math. 194 (2005), no. 1, 37–132.

[38] Pesin, Ja. B. *Characteristic Ljapunov exponents, and smooth ergodic theory.* (Russian) Uspehi Mat. Nauk 32 (1977), no. 4 (196), 55–112, 287.

[39] Katok, A. *Bernoulli diffeomorphisms on surfaces.* Ann. of Math. (2) 110 (1979), no. 3, 529–547.

[40] Dolgopyat, Dmitry; Pesin, Yakov *Every compact manifold carries a completely hyperbolic diffeomorphism.* Ergodic Theory Dynam. Systems 22 (2002), no. 2, 409–435.

[41] Shub, Michael; Wilkinson, Amie *Pathological foliations and removable zero exponents.* Invent. Math. 139 (2000), no. 3, 495–508.

[42] Pugh, Charles; Shub, Michael *Ergodic attractors.* Trans. Amer. Math. Soc. 312 (1989), no. 1, 1–54.

[43] Viana, Marcelo *Multidimensional nonhyperbolic attractors.* Inst. Hautes Études Sci. Publ. Math. No. 85 (1997), 63–96.

[44] Young, Lai-Sang *Statistical properties of dynamical systems with some hyperbolicity.* Ann. of Math. (2) 147 (1998), no. 3, 585–650.

[45] Pesin, Ja. B. *Characteristic Ljapunov exponents, and ergodic properties of smooth dynamical systems with invariant measure.* (Russian) Dokl. Akad. Nauk SSSR 226 (1976).

[46] Ruelle, David *An inequality for the entropy of differentiable maps.* Bol. Soc. Brasil. Mat. 9 (1978), no. 1, 83–87.

[47] Ledrappier, F.; Young, L.-S. *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula.* Ann. of Math. (2) 122 (1985), no. 3, 509–539.

[48] Ledrappier, F.; Young, L.-S. *The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension.* Ann. of Math. (2) 122 (1985), no. 3, 540–574.

[49] Barreira, Luis; Pesin, Yakov; Schmeling, Jorg *Dimension and product structure of hyperbolic measures.* Ann. of Math. (2) 149 (1999), no. 3, 755–783.

[50] Szász, D. *Boltzmann's ergodic hypothesis, a conjecture for centuries?* Hard ball systems and the Lorentz gas, 421–448, Encyclopaedia Math. Sci., 101, Springer, Berlin, 2000.

[51] Bunimovič, L. A. *The ergodic properties of certain billiards.* (Russian) Funkcional. Anal. i Priložen. 8 (1974), no. 3, 73–74.

[52] Chernov, Nikolai; Markarian, Roberto *Introduction to the ergodic theory of chaotic billiards.* Second edition. Publicações Matemticas do IMPA. [IMPA Mathematical Publications] 24º Colq́uio Brasileiro de Matemática. [26th Brazilian Mathematics Colloquium] Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2003.

[53] Araújo, Vitor; Pacifico, Maria José *Three Dimensional Flows.* Publicações Matemáticas do IMPA. [IMPA Mathematical Publications] 26º Colq́uio Brasileiro de Matemática. [26th Brazilian Mathematics Colloquium] Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2007.

[54] Katok, Anatole; Strelcyn, Jean-Marie; Ledrappier, F.; Przytycki, F. *Invariant manifolds, entropy and billiards; smooth maps with singularities.* Lecture Notes in Mathematics, 1222. Springer-Verlag, Berlin, 1986.

[55] Jakobson, M. V. *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps.* Comm. Math. Phys. 81 (1981), no. 1, 39–88.

[56] Graczyk, Jacek; Światek, Grzegorz *Generic hyperbolicity in the logistic family.* Ann. of Math. (2) 146 (1997), no. 1, 1–52.

[57] Lyubich, Mikhail *Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture.* Ann. of Math. (2) 149 (1999), no. 2, 319–420.

[58] Benedicks, Michael; Carleson, Lennart *The dynamics of the Hénon map.* Ann. of Math. (2) 133 (1991), no. 1, 73–169.

[59] Mora, Leonardo; Viana, Marcelo *Abundance of strange attractors.* Acta Math. 171 (1993), no. 1, 1–71.

[60] Benedicks, Michael; Young, Lai-Sang *Sinaĭ-Bowen-Ruelle measures for certain Hénon maps.* Invent. Math. 112 (1993), no. 3, 541–576.
[61] Benedicks, Michael; Viana, Marcelo *Solution of the basin problem for Hénon-like attractors.* Invent. Math. 143 (2001), no. 2, 375–434.
[62] Wang, Q.D.; Young, L.-S. *Toward a theory of rank one attractors.* Ann. of Math., to appear.
[63] Viana, M.; Lutstsatto, S. *Exclusions of parameter values in Hénon-type systems.* (Russian) Uspekhi Mat. Nauk 58 (2003), no. 6(354), 3–44; translation in Russian Math. Surveys 58 (2003), no. 6, 1053–1092
[64] Bochi, Jairo; Viana, Marcelo *Lyapunov exponents: how frequently are dynamical systems hyperbolic?* Modern dynamical systems and applications, 271–297, Cambridge Univ. Press, Cambridge, 2004.
[65] Palis, J. *A global perspective for non-conservative dynamics.* Ann. Inst. H. Poincar Anal. Non Linaire 22 (2005), no. 4, 485–507.

Northwestern University, Evanston, Illinois, USA.