Introduction

The goal of this book is to present smooth ergodic theory from a contemporary point of view. Among other things this theory provides a rigorous mathematical foundation for the phenomenon known as *deterministic chaos* – a term coined by Yorke – the appearance of highly irregular, unpredictable,"chaotic" motions in pure deterministic dynamical systems. The main idea beyond this phenomenon is that one can deduce a sufficiently complete description of topological and ergodic properties of the system from relatively weak requirements on its local behavior, known as *nonuniform hyperbolicity conditions*: the reason this theory is also called nonuniform hyperbolicity theory.

It originated in the seminal works of Lyapunov [134] and Perron [164] on stability of solutions of ordinary differential equations. To determine whether a given solution is stable one proceeds as follows. First, the equation is linearized along the solution and then the stability of the zero solution of the corresponding nonautonomous linear differential equation is examined. There are several methods (due to Hadamard [79], Perron [165], Fenichel [70], and Irwin [92]) aimed at exhibiting stability of solutions via certain information on the linear system. The approach by Lyapunov uses a special real-valued function on the space of solutions of the linear system known as the *Lyapunov exponent*. It measures in the logarithmic scale the rate of convergence of solutions so that the zero solution is asymptotically exponentially stable along any subspace where the Lyapunov exponent is negative.

The Lyapunov exponent is arguable the best way to characterize stability: the requirement that the Lyapunov exponent is negative is the weakest one that still guarantees that solutions of the linear system eventually decay exponentially to zero. The price to pay is that stability of the zero solution in this weak sense does not necessarily imply stability of the original solution of the nonlinear equation. The latter can be ensured under an additional and quite subtle requirement known as the *Lyapunov–Perron regularity*.

Verifying this requirement for a given solution may be a very difficult if not virtually impossible task, making verification more a principle than practical matter.

This could deem the whole approach useless if not for an important particular case when the differential equation is given by a vector field on a smooth compact Riemannian manifold. In this case, the celebrated Multiplicative Ergodic Theorem, also known as Oseledets' theorem, claims that a "typical" solution of the equation is Lyapunov–Perron regular, thus making the difficult task of checking the regularity requirement unnecessary. Here "typical" means that the statement holds for almost every trajectory with respect to a finite Borel measure invariant under the flow generated by the vector field.

A principal application of Oseledets' theorem in the context of smooth dynamical systems is that the Lyapunov exponent alone can be used to characterize stability of trajectories. Building upon this idea, in the beginning of 1970s Pesin introduced the class of systems whose Lyapunov exponent is nonzero along almost every trajectory with respect to some *smooth invariant measure* (i.e., a measure, which is equivalent to the Riemannian volume) and then he developed the stability theory (constructing local and global stable and unstable manifolds; see Section 7.5), as well as described their ergodic properties (including ergodicity, K- and Bernoulli properties; see Chapter 9). The collection of these results is known as Pesin's theory (see [18]). A crucial manifestation of this theory is the *formula for the entropy* connecting the measure-theoretic entropy of the system with its Lyapunov exponent (see Chapter 10). It should be pointed out that these results require that the system is of class of smoothness $C^{1+\alpha}$ for some $\alpha > 0$ and that they may indeed fail if the system is only of class C^1 (see Section 7.8).

Unlike classical uniformly hyperbolic systems (i.e., Anosov or more general axiom A systems) where contractions and expansions are *uniform everywhere* on a compact invariant set, Pesin's theory deals with systems satisfying the substantially weaker requirement that contractions and expansions occur *asymptotically almost everywhere* with respect to a smooth invariant measure. Because this requirement is weak, there are no topological obstructions for the existence of such systems on any phase space. Indeed, any smooth compact Riemannian manifold (of dimension ≥ 2 in the discrete-time case and of dimension ≥ 3 in the continuous-time case) admits a volume preserving system whose Lyapunov exponent is nonzero almost everywhere (see Sections 11.4 and 11.5). It is therefore remarkable that such a weak requirement ensures highly nontrivial ergodic and topological properties of the system.

A small perturbation of a diffeomorphism with nonzero Lyapunov exponents (in the C^r topology, $r > 1$) may not bear the same properties – the price to pay for the great generality of the nonuniformly hyperbolic theory. However, experts believe that nonuniformly hyperbolic *conservative* systems (i.e., systems preserving a smooth measure, in particular, volume preserving) are *typical* in some sense. This is reflected in the following conjectures: (We consider the case of systems with discrete time.)

1. Let f be a C^r, $r > 1$, volume preserving diffeomorphism of a smooth compact Riemannian manifold M. Assume that the Lyapunov exponent of f is nonzero along almost every trajectory of f. Then there exists a neighborhood $\mathcal{U}$

of f in the space of C^r volume preserving diffeomorphisms of M and a residual subset $\mathcal{A} \subset \mathcal{U}$ such that for every $g \in \mathcal{A}$ the Lyapunov exponent of g is nonzero along every orbit in a subset of positive volume.

2. Let f be a C^r, $r > 1$, volume preserving diffeomorphism of a smooth compact Riemannian manifold M. Then arbitrarily close to f in the space of C^r volume preserving diffeomorphisms of M, there exists a diffeomorphism g whose Lyapunov exponent is nonzero along every orbit in a subset of positive volume.

We stress that the assumption $r > 1$ is crucial as the conjectures fail if $r = 1$ due to a recent result of Bochi and Viana [28]. So far there has been little progress in solving these conjectures (see Section 11.7). On the positive side, crucial results on genericity of hyperbolic cocycles over dynamical systems have been recently obtained by Viana [221].

A persistent obstruction to nonuniform hyperbolicity is presence of elliptic behavior (see [232, 233]). For example, for area preserving surface diffeomorphisms, as predicted by KAM theory, elliptic islands survive under small perturbations of the system. Numerical studies of such maps suggest that in this case elliptic islands coexist with what appears to be a "chaotic sea" – an ergodic component of positive area with nonzero Lyapunov exponents (see [135, 136]). In fact, one often considers a one-parameter family of area preserving surface diffeomorphisms, which starts from a completely integrable (nonchaotic) system and evolves eventually into a completely hyperbolic (chaotic) one demonstrating, for intermediate values of the parameter, the appearance of elliptic islands gradually giving way to a "chaotic sea". For billiard dynamical systems, coexistence of elliptic and hyperbolic behavior has been shown for the so-called "mushroom billiards" (see [41]). In the category of smooth maps, establishing coexistence is arguably one of the most difficult problems in the theory of dynamical systems. A simple but somewhat "artificial" example of coexistence was constructed in [183] (see also [130] and Section 6.6; for a more elaborate construction see [90]). Much more complicated examples where coexistence is expected are (1) the famous standard map (also known as the Chirikov–Taylor map; see [51] and [188, Section 8.5]) and (2) automorphisms of real $K3$ surfaces (see [151]).

The requirement that the Lyapunov exponent is nonzero along almost every trajectory with respect to an invariant Borel probability measure – such a measure is said to be *hyperbolic* – is equivalent to the fact that the system is nonuniformly hyperbolic. Thus nonuniform hyperbolicity can be viewed as presence of hyperbolic invariant measures leading to challenging problems of studying ergodic and topological properties of general (not necessarily smooth) hyperbolic measures as well as of constructing some *natural* hyperbolic measures.

A general hyperbolic measure does not have "good" ergodic properties. (Simply note that *any* invariant measure on a horseshoe is hyperbolic.) It is therefore quite remarkable that hyperbolic measures have abundance of topological properties whose study was initiated in the work of Katok [101] (see Chapters 14 and 15). For example, the set of hyperbolic periodic orbits is dense in the support of the measure. Surprisingly, general hyperbolic measures asymptotically have local product structure (similar to the one of Gibbs measures on horseshoes) and one can compute their Hausdorff dimension and entropy. The formula for the entropy of a general hyperbolic measure due to Ledrappier and L.-S. Young is a substantial generalization of the entropy formula for smooth hyperbolic measures but unlike the latter, it involves quite subtle characteristics of the measure other than the Lyapunov exponent.

Smooth measures form an important yet particular case of natural hyperbolic measures. The latter were introduced by Ledrappier as an extension to nonuniformly hyperbolic systems of the Sinai–Ruelle–Bowen (SRB) measures for classical uniformly hyperbolic attractors (see Chapter 13). These *generalized* SRB measures describe the limit distribution of the time averages of continuous functions along forward orbits for a set of initial points of positive Lebesgue measure in a small neighborhood of the attractor. According to a result by Ledrappier and Strelcyn, these measures can be characterized as being the only measures for which the entropy formula of Pesin holds. Ledrappier showed that the methods used in studying ergodic properties of smooth hyperbolic measures can be adjusted to describe ergodic properties of SRB measures.

Constructing SRB measures for nonuniformly hyperbolic systems is a difficult problem. Beyond uniform hyperbolicity, there are very few examples, of which best known are Hénon-like attractors, where existence of SRB measures was rigorously shown. L.-S. Young has introduced a class of dynamical systems with nonzero Lyapunov exponents, which admit the so-called Young's tower. For these systems, she established existence of SRB measures (see Section 13.3).

The recent theory of Hénon-like diffeomorphisms (see [25, 26, 219, 222, 223]) suggests the following approach to the genericity problem for nonuniformly hyperbolic *dissipative* systems: given a one-parameter family of C^2 diffeomorphisms f_a, $a \in [\alpha, \beta]$ with a trapping region R (i.e., R is an open set for which $\overline{f_a(R)} \subset R$ for any $a \in [\alpha, \beta]$), there exists a set $A \subset [\alpha, \beta]$ of positive Lebesgue measure such that for every $a \in A$, the diffeomorphism f_a possesses an SRB measure supported on the attractor $\Lambda_a = \bigcap_{n>0} f_a^n(R)$.

Evaluating Lyapunov exponents by a computer is a relatively easy procedure and in many models in science, the absence of zero exponents can be shown numerically. This is often viewed as a convincing evidence that the system under investigation exhibits chaotic behavior. In mathematics, several "artificial" examples of systems with nonzero exponents have been constructed (and the reader can find most of them in Chapter 6) and for some interesting "natural" dynamical systems (e.g., geodesic flows on nonpositively curved manifolds and Teichmüller geodesic flows; see Chapter 12) absence of zero exponents have been shown. In addition, various powerful methods have been developed (e.g., cone and Lyapunov function techniques; see Chapter 4) that allow one to verify whether a given dynamical system has some positive Lyapunov exponents.

Many results of the nonuniform hyperbolicity theory hold in greater generality than for actions of single dynamical systems and wherever possible we describe the theory with this view in mind. For example, the linear hyperbolicity theory (including the theory of Lyapunov exponents and its principal result – the Multiplicative Ergodic Theorem) is presented for linear cocycles over dynamical systems (or even over higher-rank Abelian actions), and the stable manifold theory is developed for sequences of diffeomorphisms. Even in the case of an action of a single dynamical system, we consider a more general case of nonuniform *partial* hyperbolicity where the requirement that the values of the Lyapunov exponent are *all* nonzero is replaced by a weaker one that *some* of the values of the Lyapunov exponent are nonzero. Such generalizations require some more complicated techniques and tools from various areas of mathematics to be used and thus make the exposition more complicated but they substantially broaden applications and show the great power of the nonuniform hyperbolicity theory.

Cambridge University Press
978-0-521-83258-8 - Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents
Luis Barreira and Yakov Pesin
Excerpt
More information

Part I

Linear Theory

1

The Concept of Nonuniform Hyperbolicity

In this chapter we consider sequences of linear maps in the Euclidean space and we introduce the principal notions of Lyapunov exponents, Lyapunov–Perron regularity, normal bases, and so on. These concepts are used in a variety of settings of which the main one is the study of linearizations of a dynamical system along its orbits. Thus a sequence of linear maps can be thought of as the sequence of derivatives (differentials) of a smooth map along an orbit.

We stress that in the situations we consider there are no preferred coordinate systems. Accordingly, even though we often use matrix representations of linear maps, we only study properties that are independent of certain classes of coordinate changes. The most narrow class is that of orthogonal coordinate changes; in the smooth situation, this corresponds to fixing a Riemannian metric in the phase space. A broader class includes coordinate changes uniformly bounded from above and below; in the case of a smooth system on a compact space, this corresponds to an arbitrary choice of a smooth coordinate atlas.

As it turns out, of greatest importance for the theory developed in this book is still a broader class of tempered coordinate changes. This reflects the primary role that exponential behavior plays in our considerations. A tempered change allows arbitrarily large distortions if these distortions change with time with a subexponential rate. Thus it preserves not only the exponential character of the asymptotic behavior but also the actual rates of expansion and contraction in various directions.

In the next chapter these considerations will be extended to the case of cocycles over dynamical systems. The principal difference is in paying attention to the dependence on the base point that may be measurable, continuous, differentiable, and so on.

1.1 Motivation

Consider an invertible linear transformation A of an Euclidean space $\mathbb{R}^n$. Let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues. The transformation A is said to be *hyperbolic* if

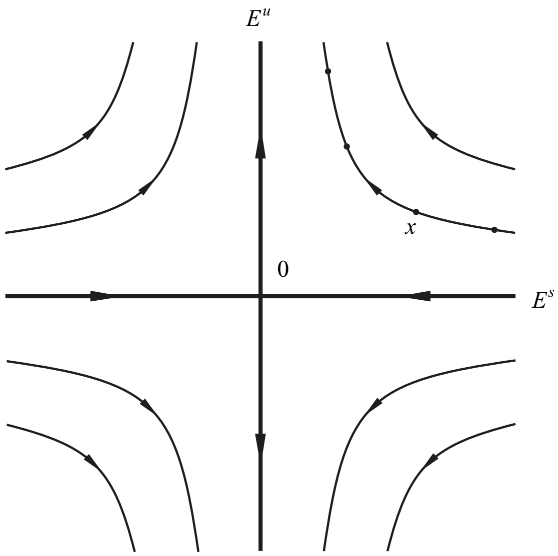

Figure 1.1. Hyperbolic fixed point of a linear map: the trajectory of a point x lies on the hyperbola passing through x.

$|\lambda_i| \neq 1$ for every i. If A is hyperbolic then there is a decomposition $\mathbb{R}^n = E^s \oplus E^u$ into *stable* and *unstable subspaces*, where

$$E^s = \mathbb{R}^n \cap \bigoplus_{i:|\lambda_i|<1} H_{\lambda_i} \quad \text{and} \quad E^u = \mathbb{R}^n \cap \bigoplus_{i:|\lambda_i|>1} H_{\lambda_i}.$$

Here H_{λ_i} is the root space

$$H_{\lambda_i} = \{v \in \mathbb{C}^n : (A - \lambda_i \mathrm{Id})^m v = 0 \text{ for some } m \in \mathbb{N}\}.$$

Set

$$\lambda = \max\left\{\max_{|\lambda_i|<1} |\lambda_i|, \max_{|\lambda_i|>1} |\lambda_i|^{-1}\right\} \in (0, 1).$$

Note that for any $\varepsilon > 0$, there exists $c = c(\varepsilon) > 0$ such that for every $m > 0$,

$$\|A^m v\| \leq c\lambda^m e^{\varepsilon m} \|v\| \quad \text{if } v \in E^s,$$

and

$$\|A^{-m} v\| \leq c\lambda^m e^{\varepsilon m} \|v\| \quad \text{if } v \in E^u.$$

Therefore the origin is a saddlelike point for A (see Fig. 1.1).

Notice that if A is diagonalizable, a similar estimate holds with $\delta = 0$ but if A has Jordan blocks then the growth in the corresponding root space has an extra subexponential factor.

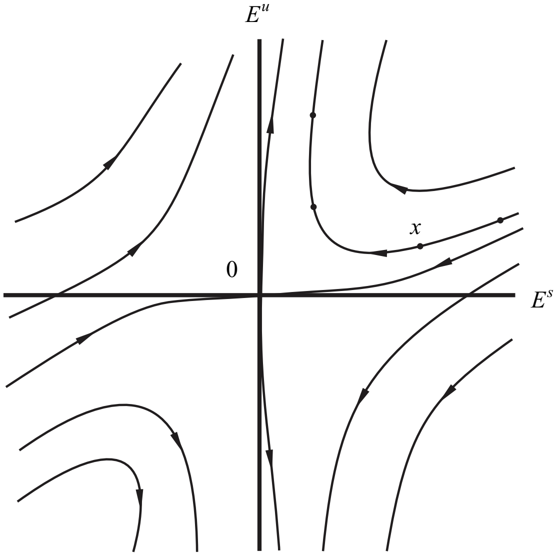

Figure 1.2. Hyperbolic fixed point of a diffeomorphism: near 0 the trajectory of a point x lies on a curve, which is the pre-image of a hyperbola under the conjugacy map h.

Let now f be a C^1 diffeomorphism of an open set U of $\mathbb{R}^n$ and $p \in U$ a *fixed point* for f, that is, $f(p) = p$. The point p is called *hyperbolic* if the linear transformation $A = d_p f$ is hyperbolic. By the Grobman–Hartman Theorem (see e.g., [104, Theorem 6.3.1]), there is a homeomorphism h defined in a small neighborhood $\widetilde{U} \subset U$ of p such that $h(f(x)) = A(h(x))$ for every $x \in \widetilde{U} \cap f^{-1}(\widetilde{U})$. This implies that the local orbit structure of a diffeomorphism in a small neighborhood of a hyperbolic fixed point resembles that of the linear transformation that is the differential (the linearization) of the map at that point (see Fig. 1.2).

By the Hadamard–Perron Theorem (see e.g., [104, Theorem 6.4.9]), the *stable set*

$$V^s(p) = \{x \in U : f^m(x) \in U \text{ for all } m > 0\}$$

and the *unstable set*

$$V^u(p) = \{x \in U : f^m(x) \in U \text{ for all } m < 0\}$$

are C^1 submanifolds passing through p such that $T_p V^s(p) = E^s$ and $T_p V^u(p) = E^u$. Moreover, similarly to the linear case above, given $\varepsilon > 0$, there exists $c = c(\varepsilon) > 0$ such that for every $m > 0$,

$$\|f^m(x) - p\| \le c\lambda^m e^{\varepsilon m} \|x\| \quad \text{if } x \in V^s(p),$$

and

$$\|f^{-m}(x) - p\| \le c\lambda^m e^{\varepsilon m}\|x\| \quad \text{if } x \in V^u(p).$$

We call $V^s(p)$ and $V^u(p)$ the *local stable* and *unstable manifolds* at p.

This discussion extends directly to hyperbolic periodic points. A point $p \in U$ is *periodic* if for some $k \in \mathbb{N}$, we have $f^i(p) \in U$ for $i = 1, \dots, k-1$ and $f^k(p) = p$. The periodic point p is called *hyperbolic* if the linear transformation $d_p f^k$ is hyperbolic. One can show that there exist open neighborhoods $U_0, \dots, U_{k-1}$, respectively, of $p, \dots, f^{k-1}(p)$ such that for each $i = 0, \dots, k-1$, the sets

$$V^s(f^i(p)) = \{x \in U_i : f^m(x) \in U_{m+i} \text{ for all } m > 0\}$$

and

$$V^u(f^i(p)) = \{x \in U_i : f^m(x) \in U_{m+i} \text{ for all } m < 0\}$$

are C^1 submanifolds passing through $f^i(p)$. We call $V^s(f^i(p))$ and $V^u(f^i(p))$ the *local stable* and *unstable manifolds* at $f^i(p)$. Let us point out that if p is a hyperbolic periodic point of period k then for every $i = 1, \dots, k-1$, the point $f^i(p)$ is a hyperbolic periodic point. In other words, the trajectory $\{f^i(p) : i = 0, \dots, k-1\}$ is hyperbolic.

Now consider a nonperiodic point p of a diffeomorphism f. We would like to say that p is hyperbolic if the behavior of trajectories that start in a neighborhood of p resembles that of the trajectories in a neighborhood of a periodic hyperbolic point. More precisely, this means that one can construct local stable and unstable manifolds $V^s(p)$ and $V^u(p)$ at p such that every trajectory $f^m(v)$ with $v \in V^s(p)$ approaches $f^m(p)$ with an exponential rate and every trajectory $f^{-m}(v)$ with $v \in V^u(p)$ approaches $f^{-m}(p)$ with an exponential rate. In fact the orbit $\{f^m(x)\}_{m\in\mathbb{Z}}$ of a point $x \in V^s(p)$ may first diverge from the orbit $\{f^m(p)\}_{m\in\mathbb{Z}}$ until some time $T = T(x)$ before it starts converging to it with an exponential rate. If the rate of the exponential convergence does not depend on the point p in a compact f-invariant subset $\Lambda \subset U$, and the angle between local stable and unstable manifolds at p is uniformly bounded away from zero in p, then f is said to be *uniformly hyperbolic* on Λ. In this case the function $T(x)$ is uniformly bounded from above in x. Note that hyperbolicity refers to the whole orbit $\{f^m(p)\}_{m\in\mathbb{Z}}$.

In the general case, however, the rate of exponential convergence may vary from orbit to orbit. Moreover, the function $T(x)$ may not be bounded from above as x approaches p, forcing contraction estimates to deteriorate along the orbit. If this deterioration occurs with a subexponential rate (or a sufficiently small exponential rate) on a set Λ of orbits, then f is said to be *nonuniformly hyperbolic* on Λ. Observe that this set is invariant but not necessarily closed. The dynamics of f on the nonuniformly hyperbolic set is the main object of our study.

In view of the Grobman–Hartman Theorem, it is natural to try to introduce hyperbolicity in terms of the sequence of differentials $d_p f^m$. This leads us to the consideration of a sequence of linear transformations of an Euclidean space. It is important to keep in mind that unlike the case of a fixed or periodic point,