Lyapunov Exponents and Smooth Ergodic Theory

Luis Barreira and Yakov B. Pesin

2000 *Mathematics Subject Classification.* Primary: 37D25, 37C40.

Abstract. This book provides a systematic introduction to smooth ergodic theory, including the general theory of Lyapunov exponents, nonuniform hyperbolic theory, stable manifold theory emphasizing absolute continuity of invariant foliations, and the ergodic theory of dynamical systems with nonzero Lyapunov exponents. The book can be used as a primary textbook for a special topics course on nonuniform hyperbolicity or as supplementary reading for a basic course on dynamical systems.

Contents

Preface

This book provides a systematic introduction to the core of smooth ergodic theory. Despite an impressive amount of literature in the field there is no textbook which contains a sufficiently complete presentation of the theory. This book attempts to fill this gap. We describe the general (abstract) theory of Lyapunov exponents and its applications to the stability theory of differential equations, the stable manifold theory, absolute continuity of stable manifolds, and the ergodic theory of dynamical systems with nonzero Lyapunov exponents (including geodesic flows).

The book is a revised and considerably expanded version of our *Lectures on Lyapunov Exponents and Smooth Ergodic Theory* [**4**]. We added more examples of dynamical systems with nonzero Lyapunov exponents, including diffeomorphisms on two-dimensional tori and on spheres. Furthermore, we substantially expanded the exposition of the crucial absolute continuity property. In particular, we include an example of a foliation that is not absolutely continuous and establish the formula for the Jacobian of the holonomy map. We also added a complete proof of the Multiplicative Ergodic Theorem as well as provided more details in the proofs of several basic results. Finally, a few more figures were added to illustrate the exposition.

We hope that these improvements make the book more accessible to graduate students or anyone who wishes to acquire a working knowledge of smooth ergodic theory and to learn how to use its tools. Indeed, the book can be used as a primary textbook for a special topics course on nonuniform hyperbolic theory or as supplementary reading for a basic course on dynamical systems.

This book is self-contained. We only assume that the reader has a basic knowledge of real analysis, measure theory, differential equations, and topology. We present the basic concepts of smooth ergodic theory and provide complete proof of all main results. We also state some results whose proofs require more advanced techniques which exceed the scope of the book. In our opinion this gives the reader a broader view of smooth ergodic theory and may help stimulate further study. This will also provide nonexperts with a broader perspective of the field.

While writing this book we consulted with Anatole Katok on several topics and we would like to thank him for his many valuable comments.

We also would like to thank Misha Brin, Boris Hasselblatt, Jörg Schmeling, and Howie Weiss for many useful remarks on mathematical structure, style, references, etc. It is also our pleasure to thank Ilie Ugarcovici and Alistair Windsor, graduate students at Penn State, who read the text thoroughly and helped us correct some typos and mistakes and improve the exposition. We especially thank Natasha Pesin, an experienced editor, for her editorial assistance.

August 2001

Luis Barreira
Lisboa, Portugal

Yakov B. Pesin
State College, PA, USA

Introduction

Smooth ergodic theory studies the ergodic properties of smooth dynamical systems on Riemannian manifolds with respect to "natural" invariant measures. The most important of such measures are smooth measures, i.e., measures that are equivalent to the Riemannian volume. There are various classes of smooth dynamical systems whose study requires different techniques. In this book we concentrate on systems whose trajectories are hyperbolic in some sense. Roughly speaking, this means that the behavior of trajectories near a given orbit resembles the behavior of trajectories near a saddle point. In particular, to every hyperbolic trajectory one can associate two complementary subspaces such that the system acts as a contraction along one of them (called the stable subspace) and as an expansion along the other (called the unstable subspace).

A hyperbolic trajectory is unstable – almost every nearby trajectory moves away from it with time. If the set of hyperbolic trajectories is sufficiently large (for example, has positive or full measure), this instability forces trajectories to become separated. If the phase space of the system is compact, the trajectories mix together because there is not enough room to separate them. This is one of the main reasons why systems with hyperbolic trajectories on compact phase spaces exhibit chaotic behavior. Indeed, hyperbolic theory provides a mathematical foundation for the paradigm that is widely known as "deterministic chaos" – the appearance of irregular chaotic motions in purely deterministic dynamical systems. This paradigm asserts that conclusions about global properties of a nonlinear dynamical system with sufficiently strong hyperbolic behavior can be deduced from studying the linearized systems along its trajectories.

The study of hyperbolic phenomena originated in seminal works of Artin, Morse, Hedlund, and Hopf on the instability and ergodic properties of geodesic flows on compact surfaces (see the survey [**18**] for a detailed description of results obtained at this time and for references). Later, hyperbolic behavior was observed in other situations (for example, Smale horseshoes and hyperbolic toral automorphism). The systematic study of hyperbolic dynamical systems was initiated by Smale (who mainly considered the problem of structural stability of hyperbolic systems; see [**40**]) and by Anosov and Sinai (who were mainly concerned with ergodic properties of hyperbolic systems with respect to smooth invariant measures; see [**2, 3**]). The hyperbolicity conditions describe the action of the linearized system along the stable and unstable subspaces and impose quite strong requirements on the system. The dynamical systems that satisfy these hyperbolicity conditions uniformly over all orbits are called Anosov systems. They possess strong ergodic properties which we discuss in Chapter 5.

In this book we consider the weakest (hence, most general) form of hyperbolicity known as nonuniform hyperbolicity. This was introduced by Pesin in [**31, 32, 33, 34**] and its study (which is sometimes referred to as "Pesin theory") is based upon the theory of Lyapunov exponents. The latter originated in works of Lyapunov [**26**] and Perron [**30**] and was developed further in [**11**]. We provide an extended excursion into the theory of Lyapunov exponents and in particular, introduce and study the crucial concept of Lyapunov–Perron regularity (see Sections 1.3 and 1.5). The theory of Lyapunov exponents enables one to obtain many subtle results on stability of differential equations (see Section 1.4). We stress that many of these results cannot be found in the English mathematical literature.

In Chapter 2 we adapt the main results of the abstract theory of Lyapunov exponents as well as the concept of Lyapunov–Perron regularity to dynamical systems. The reader who is primarily interested in studying nonuniform hyperbolic theory can skip Chapter 1 and proceed directly to Chapter 2. However, we would like to stress that the proofs of the results on stability of differential equations contain ,in a simple form, many ideas which are used later to study stability of nonuniformly hyperbolic systems.

Using the language of Lyapunov exponents one can view nonuniformly hyperbolic dynamical systems as those systems where the set of points with *all* Lyapunov exponents nonzero is "large" – for example, has full measure with respect to an invariant Borel measure (see Sections 2.1 and 2.2). In this case the Multiplicative Ergodic theorem of Oseledets [**29**] implies that almost every point is Lyapunov–Perron regular (see Sections 2.2 and 2.4). The powerful theory of Lyapunov exponents then yields a profound description of the local stability of trajectories. This is the subject of study in Chapter 4.

The crucial difference between the classical uniform hyperbolicity and its weakened version of nonuniform hyperbolicity is that the hyperbolicity conditions may deteriorate when one moves along a nonuniformly hyperbolic trajectory. However, if the trajectory is Lyapunov–Perron regular then the deterioration occurs at a subexponential rate and the contraction and expansion along stable and unstable directions prevail. A crucial application of this fact is the Stable Manifold Theorem, which was established by Pesin in [**32**], and is a substantial generalization of the classical Hadamard–Perron theorem.

In Section 4.1 we present the proof of the Stable Manifold theorem following the original approach in [**32**] which is essentially an elaboration of the Perron method. In Section 4.2 we sketch the proof of a slightly more general version of the Stable Manifold theorem (known as the Graph Transform Property) that is due to Hadamard. Since hyperbolicity is nonuniform, the sizes of local stable manifolds may decrease along trajectories. However, due to Lyapunov–Perron regularity this can only occur at subexponential rate.

There are several methods for establishing nonuniform hyperbolicity. One of them, which we consider in this book, is to show that the Lyapunov exponents of the system are nonzero. The first example of a system, which has nonzero Lyapunov exponents but is not an Anosov system, was constructed in [**31**]. It is a three dimensional flow on a compact smooth Riemannian manifold and is a modification of an Anosov flow. The example reveals some general mechanisms which cause zero Lyapunov exponents. We elaborate on such mechanisms in Section 3.3.

Another example of a system with nonzero Lyapunov exponents is the geodesic flow on a compact smooth Riemannian manifold of nonpositive curvature. Geodesic flows have always been a rich source of examples and have provided inspiration for developing hyperbolic theory. For instance, the study of geodesic flows on compact Riemannian manifolds of negative curvature stimulated the development of uniform hyperbolic theory while geodesic flows on Riemannian manifolds of nonpositive curvature were a challenging application of nonuniform hyperbolic theory. In this book we consider only geodesic flows on surfaces of nonpositive curvature and show that they are nonuniformly hyperbolic on an open and dense set (see Section 3.4).

Hyperbolic theory supplies ideas and methods which enable one to study stochastic properties of hyperbolic systems with respect to smooth invariant measures. This study includes describing ergodic components and computing the measure theoretical entropy. These are the main topics of Chapter 5.

The description of ergodic components begins with a beautiful argument due to Hopf that allows one to study ergodicity of the system by analyzing the local behavior of its trajectories. This argument exploits a crucial property of local manifolds known as absolute continuity. Roughly speaking, this means that one can apply the Fubini theorem to the partition of the phase space formed by the local manifolds. We discuss the absolute continuity property in Sections 4.3, 4.4, and 5.1.

In Section 5.2 we show that a nonuniformly hyperbolic system has ergodic components of positive measure. In Section 5.4 we introduce the reader to the local ergodicity problem, i.e., find conditions which guarantee that ergodic components are open (mod 0). In Section 5.4 we establish the entropy formula that expresses the entropy of the system via its positive Lyapunov exponents. Finally, in Section 5.6 we apply these results to geodesic flows on compact surfaces of nonpositive curvature.

Using more sophisticated ideas and more subtle techniques, one can obtain refined information on the ergodic properties of nonuniformly hyperbolic dynamical systems with respect to smooth measures, and in particular, establish the Bernoulli property. Moreover, one can extend many of the above mentioned results to some other classes of invariant measures such as Sinai–Ruelle–Bowen measures. Although the detailed description of these results is beyond the scope of this book, we include a brief discussion of the Bernoulli property (see Section 5.2), the local product structure and the dimension of general hyperbolic measures, as well as provide a characterization of Sinai–Ruelle–Bowen measures (see Section 5.5).