Foundations of Ergodic Theory

Marcelo Viana and Krerley Oliveira

Preface

In short terms, Ergodic Theory is the mathematical discipline that deals with dynamical systems endowed with invariant measures. Let us begin by explaining what we mean by this and why these mathematical objects are so worth studying. Next, we highlight some of the major achievements in this field, whose roots go back to the Physics of the late 19th century. Near the end of the preface, we outline the content of this book, its structure and its pre-requisites.

What is a dynamical system?

There are several definitions of what a dynamical system is, some more general than others. We restrict ourselves to two main models.

The first one, to which we refer most of the time, is a transformation $f : M \to M$ in some space M. Heuristically, we think of M as the space of all possible states of a given system. Then f is the evolution law, associating with each state $x \in M$ the one state $f(x) \in M$ the system will be in a unit of time later. Thus, time is a discrete parameter in this model.

We also consider models of dynamical systems with continuous time, namely flows. Recall that a *flow* in a space M is a family $f^t : M \to M$, $t \in \mathbb{R}$ of transformations satisfying

$$f^0 = \text{ identity} \quad \text{and} \quad f^t \circ f^s = f^{t+s} \text{ for all } t, s \in \mathbb{R}. \tag{0.0.1}$$

Flows appear, most notably, in connection with differential equations: take f^t to be the transformation associating with each $x \in M$ the value at time t of the solution of the equation that passes through x at time zero.

We always assume that the dynamical system is measurable, that is, that the space M carries a σ-algebra of *measurable subsets* that is preserved by the dynamics, in the sense that the pre-image of any measurable subset is still a measurable subset. Often, we take M to be a topological space, or even a metric space, endowed with the Borel σ-algebra, that is, the smallest σ-algebra that contains all open sets. Even more, in many of the situations we consider in this book, M is a smooth manifold and the dynamical system is taken to be differentiable.

What is an invariant measure?

A *measure* in M is a non-negative function μ defined on the σ-algebra of M, such that $\mu(\emptyset) = 0$ and

$$\mu\big(\cup_n A_n\big) = \sum_n \mu(A_n)$$

for any countable family $\{A_n\}$ of pairwise disjoint measurable subsets. We call μ a probability measure if $\mu(M) = 1$. In most cases, we deal with finite measures, that is, such that $\mu(M) < \infty$. Then we can easily turn μ into a probability ν: just define

$$\nu(E) = \frac{\mu(E)}{\mu(M)} \quad \text{for every measurable set } E \subset M.$$

In general, we say that a measure μ is invariant under a transformation f if

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{for every measurable set } E \subset M. \tag{0.0.2}$$

Heuristically, this may be read as follows: the probability that a point is in any given measurable set is the same as the probability that its image is in that set. For flows, we replace (0.0.2) by

$$\mu(E) = \mu(f^{-t}(E)) \quad \text{for every measurable set } E \subset M \text{ and } t \in \mathbb{R}. \tag{0.0.3}$$

Notice that (0.0.2)–(0.0.3) do make sense since, by assumption, the pre-image of a measurable set is also a measurable set.

Why study invariant measures?

As in any other branch of mathematics, an important part of the motivation is intrinsic and aesthetical: as we will see, these mathematical structures have deep and surprising properties, which are expressed through beautiful theorems. Equally fascinating, ideas and results from Ergodic Theory can be applied in many other areas of Mathematics, including some that do not seem to have anything to do with probabilistic concepts, such as Combinatorics and Number Theory.

Another key motivation is that many problems in the experimental sciences, including many complicated natural phenomena, can be modelled by dynamical systems that leave some interesting measure invariant. Historically, the most important example came from Physics: Hamiltonian systems, which describe the evolution of conservative systems in Newtonian Mechanics, are described by certain flows that preserve a natural measure, the so-called Liouville measure. Actually, we will see that very general dynamical systems do possess invariant measures.

Yet another fundamental reason to be interested in invariant measures is that their study may yield important information on the dynamical system's behavior that would be difficult to obtain otherwise. Poincaré's recurrence theorem, one of the first results we analyze in this book, is a great illustration of this: it asserts that, relative to any finite invariant measure, almost every orbit returns arbitrarily close to its initial state.

Brief historic survey

The word *ergodic* is a concatenation of two Greek words, $\epsilon\rho\gamma o\nu$ (*ergon*) = work and $o\delta o\sigma$ (*odos*) = way, and was introduced in the 19th century by the Austrian physicist L. Boltzmann. The systems that interested Boltzmann, J. C. Maxwell and J. C. Gibbs, the founders of the kinetic theory of gases, can be described by a Hamiltonian flow, associated with a differential equation of the form

$$\left(\frac{dq_1}{dt}, \dots, \frac{dq_n}{dt}, \frac{dp_1}{dt}, \dots, \frac{dp_n}{dt}\right) = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n}\right).$$

Boltzmann believed that typical orbits of such a flow fill in the whole energy surface $H^{-1}(c)$ that contains them. Starting from this *ergodic hypothesis*, he deduced that the (time) averages of observable quantities along typical orbits coincide with the (space) averages of such quantities on the energy surface, which was crucial for his formulation of the kinetic theory of gases.

In fact, the way it was formulated originally by Boltzmann, this hypothesis is clearly false. So, the denomination *ergodic hypothesis* was gradually displaced to what would have been a consequence, namely, the claim that time averages and space averages coincide. Systems for which this is true were called *ergodic*. And it is fair to say that a great part of the progress experienced by Ergodic Theory in the 20th century was motivated by the quest to understand whether most Hamiltonian systems, especially those that appear in connection with the kinetic theory of gases, are ergodic or not.

The foundations were set in the 1930's, when J. von Neumann and G. D. Birkhoff proved that time averages are indeed well defined for almost every orbit. However, in the mid 1950's, the great Russian mathematician A. N. Kolmogorov observed that many Hamiltonian systems are actually *not* ergodic. This spectacular discovery was much expanded by V. Arnold and J. Moser, in what came to be called KAM (Kolmogorov-Arnold-Moser) theory.

On the other hand, still in the 1930's, E. Hopf had given the first important examples of Hamiltonian systems that *are* ergodic, namely, the geodesic flows on surfaces with negative curvature. His result was generalized to geodesic flows on manifolds of any dimension by D. Anosov, in the 1960's. In fact, Anosov proved ergodicity for a much more general class of systems, both with discrete time and in continuous time, which are now called Anosov systems.

An even broader class, called uniformly hyperbolic systems, was introduced by S. Smale and became a major focus for the theory of Dynamical Systems through the last half a century or so. In the 1970's, Ya. Sinai developed the theory of Gibbs measures for Anosov systems, conservative or dissipative, which D. Ruelle and R. Bowen rapidly extended to uniformly hyperbolic systems. This certainly ranks among the greatest achievements of smooth ergodic theory.

Two other major contributions must also be mentioned in this brief survey. One is the introduction of the notion of *entropy*, by Kolmogorov and Sinai, near the end of the 1950's. Another is the proof that the entropy is a complete invariant for Bernoulli shifts (two Bernoulli shifts are equivalent if and only if they have the same entropy), by D. Ornstein, some ten years later.

By then, the theory of non-uniformly hyperbolic systems was being initiated by V. I. Oseledets, Ya. Pesin and others. But that would take us beyond the scope of the present book.

How this book came to be

This book grew from lecture notes we wrote for the participants of mini-courses we taught at the Department of Mathematics of the Universidade Federal de Pernambuco (Recife, Brazil), in January 2003, and at the meeting *Novos Talentos em Matemática* held by Fundação Calouste Gulbenkian (Lisbon, Portugal), in September 2004.

In both cases, most of the audience consisted of young undergraduates with little previous contact with Measure Theory, let alone Ergodic Theory. Thus, it was necessary to provide very friendly material that allowed such students to follow the main ideas to be presented. Still at that stage, our text was used by other colleagues, such as Vanderlei Horita (São José do Rio Preto, Brazil), for teaching mini-courses to audiences with a similar profile.

As the text evolved, we have tried to preserve this elementary character of the early chapters, especially Chapters 1 and 2, so that they can used independently of the rest of the book, with as few prerequisites as possible.

Starting from the mini-course we gave at the 2005 *Colóquio Brasileiro de Matemática* (IMPA, Rio de Janeiro), this project acquired a broader purpose. Gradually, we evolved towards trying to present in a consistent textbook format the material that, in our view, constitutes the core of Ergodic Theory. Inspired by our own research experience in this area, we endeavored to assemble in a unified presentation the ideas and facts upon which is built the remarkable development this field experienced over the last decades.

A main concern was to try and keep the text as self-contained as possible. Ergodic Theory is based on several other mathematical disciplines, especially Measure Theory, Topology and Analysis. In the appendix, we have collected the main material from those disciplines that is used throughout the text. As a rule, proofs are omitted, since they can easily be found in many of the excellent references we provide. However, we do assume that the reader is familiar with the main tools of Linear Algebra, such as the canonical Jordan form.

Structure of the book

The main part of this book consists of 12 chapters, divided into sections and subsections, and one appendix, also divided into sections and subsections. A list of exercises is given at the end of every section, appendix included. Statements (theorems, propositions, lemmas, corollaries, etc.), exercises and formulas are numbered by section and chapter: for instance, (2.3.7) is the seventh formula in the third section of the second chapter and Exercise A.5.1 is the first exercise in the fifth section of the appendix. Hints for selected exercises are given in a special chapter after the appendix. At the end, we provide a list of references and an index.

Chapters 1 through 12 are organized as follows:

- Chapters 1 through 4 constitute a kind of introductory cycle, in which we present the basic notions and facts in Ergodic Theory - invariance, recurrence and ergodicity - as well as some main examples. Chapter 3 introduces the fundamental results (ergodic theorems) upon which the whole theory is built.
- Chapter 4, where we introduce the key notion of ergodicity, is a turning point in our text. The next two chapters (Chapters 5 and 6) develop a couple of important related topics: decomposition of invariant measures into ergodic measures and systems admitting a unique, necessarily ergodic, invariant measure.
- Chapters 7 through 9 deal with very diverse subjects - loss of memory, the isomorphism problem and entropy - but they also form a coherent structure, built around the idea of considering increasingly "chaotic" systems: mixing, Lebesgue spectrum, Kolmogorov and Bernoulli systems.
- Chapter 9 is another turning point. As we introduce the fundamental concept of entropy, we take our time to present it to the reader from several different viewpoints. This is naturally articulated with the content of Chapter 10, where we develop the topological version of entropy, including an important generalization called pressure.
- In the two final chapters, 11 and 12, we focus on a specific class of dynamical systems, called expanding transformations, that allows us to exhibit a concrete (and spectacular!) application of many of the general ideas presented in the text. This includes Ruelle's theorem and its applications, which we view as a natural climax of the book.

Appendices A.1 through A.2 cover several basic topics of Measure and Integration. Appendix A.3 deals with the special case of Borel measures in metric spaces. In Appendix A.4 we recall some basic facts from the theory of manifolds and smooth maps. Similarly, Appendices A.5 and A.6 cover some useful basic material about Banach spaces and Hilbert spaces. Finally, Appendix A.7 is devoted to the spectral theorem.

Examples and applications have a key part in any mathematical discipline and, perhaps, even more so in Ergodic Theory. For this reason, we devote special attention to presenting concrete situations that illustrate and put in perspective the general results. Such examples and constructions are introduced gradually, whenever the context seems better suited to highlight their relevance. They often return later in the text, to illustrate new fundamental concepts as we introduce them.

The exercises at the end of each section have a threefold purpose. There are routine exercises meant to help the reader become acquainted with the concepts and the results presented in the text. Also, we leave as exercises certain arguments and proofs that are not used in the sequel or belong to more elementary related areas, such as Topology or Measure Theory. Finally, more sophisticated exercises test the reader's global understanding of the theory. For the reader's convenience, hints for selected exercises are given in a special chapter following the appendix.

How to use this book?

These comments are meant, primarily, for the reader who plans to use this book to teach a course. Appendices A.1 through A.7 provide quick references to background material. In principle, they are not meant to be presented in class.

The content of Chapters 1 through 12 is suitable for a one-year course, or a sequence of two one-semester courses. In either case, the reader should be able to cover most of the material, possibly reserving some topics for seminars given by the students. The following sections are especially suited for that:

Section 1.5, Section 2.5, Section 3.4, Section 4.4, Section 6.4, Section 7.3, Section 7.4, Section 8.3, Section 8.4, Section 8.5, Section 9.5, Section 9.7, Section 10.4, Section 10.5, Section 11.1, Section 11.3, Section 12.3 and Section 12.4.

In this format, Ruelle's theorem (Theorem 12.1) and its applications are a natural closure for the course.

In case only one semester is available, some selection of topics will be necessary. The authors' suggestion is to try and cover the following program:

Chapter 1: Sections 1.1, 1.2 and 1.3.
Chapter 2: Sections 2.1 and 2.2.
Chapter 3: Sections 3.1, 3.2 and 3.3.
Chapter 4: Sections 4.1, 4.2 and 4.3.
Chapter 5: Section 5.1 (mention Rokhlin's theorem).
Chapter 6: Sections 6.1, 6.2 and 6.3.
Chapter 7: Sections 7.1 and 7.2.
Chapter 8: Section 8.1 and 8.2 (mention Ornstein's theorem).
Chapter 9: Sections 9.1, 9.2, 9.3 and 9.4.
Chapter 10: Sections 10.1 and 10.2.
Chapter 11: Section 11.1.

In this format, the course could close either with the proof of the variational principle for the entropy (Theorem 10.1) or with the construction of absolutely continuous invariant measures for expanding maps on manifolds (Theorem 11.1.2).

We have designed the text in such a way as to make it feasible for the lecturer to focus on presenting the central ideas, leaving it to the student to study in detail many of the proofs and complementary results. Indeed, we devoted considerable effort to making the explanations as friendly as possible, detailing the arguments and including plenty of cross-references to previous related results as well to the definitions of the relevant notions.

In addition to the regular appearance of examples, we have often chosen to approach the same notion more than once, from different points of view, if that seemed useful for its in-depth understanding. The special chapter containing the hints for selected exercises is also part of that effort to encourage and facilitate the autonomous use of this book by the student.

Acknowledgments

The writing of this book extended for over a decade. During this period we benefitted from constructive criticism from several colleagues and students.

Many colleagues used different preliminary versions of the book to teach courses and shared their experiences with us. Besides Vanderlei Horita (São José do Rio Preto, Brazil), Nivaldo Muniz (São Luis, Brazil) and Meysam Nassiri (Teheran, Iran), we would like to thank Vítor Araújo (Salvador, Brazil) for an extended list of suggestions that influenced significantly the way the text evolved from then on. François Ledrappier (Paris, France) helped us with questions about substitution systems.

We also had the chance to test the material in a number of regular graduate courses at IMPA-Instituto de Matemática Pura e Aplicada and at UFAL-Universidade Federal de Alagoas. Feedback from graduate students Adriana Sánchez, Aline Gomes Cerqueira, El Hadji Yaya Tall, Ermerson Araujo, Ignacio Atal, Rafael Lucena, Raphaël Cyna and Xiao-Chuan Liu allowed us to correct many of the weaknesses in earlier versions.

The first draft of Appendices A.1-A.2 was written by João Gouveia, Vítor Saraiva and Ricardo Andrade, who acted as assistants for the course in *Novos Talentos em Matemática 2004* mentioned previously. IMPA students Edileno de Almeida Santos, Felippe Soares Guimarães, Fernando Nera Lenarduzzi, Ítalo Dowell Lira Melo, Marco Vinicius Bahi Aymone and Renan Henrique Finder wrote many of the hints for the exercises in Chapters 1 through 8 and the appendix.

The original Portuguese version of this book, *Fundamentos da Teoria Ergódica* [VO14], was published in 2014 by SBM-Sociedade Brasileira de Matemática. Feedback from colleagues who used that back to teach graduate courses in different places helped eliminate some of the remaining shortcomings. The extended list of remarks by Bernardo Lima (Belo Horizonte, Brazil) and his student Leonardo Guerini was particularly useful in this regard.

Several other changes and corrections were made in the course of the translation to English. We are grateful to the many colleagues and students who agreed to revise different parts of the translated text, especially Cristina Lizana, Elaís C. Malheiro, Fernando Nera Lenarduzzi, Jiagang Yang, Karina Marin, Lucas Backes, Maria João Resende, Mauricio Poletti, Paulo Varandas, Ricardo Turolla, Sina Türelli, Vanessa Ramos, Vítor Araújo and Xiao-Chuan Liu.

Rio de Janeiro and Maceió, March 31, 2015

Marcelo Viana[1] and Krerley Oliveira[2]

[1]IMPA, Estrada D. Castorina 110, 22460-320 Rio de Janeiro, Brasil. viana@impa.br.
[2]Departamento de Matemática, Universidade Federal de Alagoas, Campus A. C. Simões s/n, 57072-090 Maceió, Brasil. krerley@mat.ufal.br.

Contents

Chapter 1

Recurrence

Ergodic Theory studies the behavior of dynamical systems with respect to measures that remain invariant under time evolution. Indeed, it aims to describe those properties that are valid for the trajectories of almost all initial states of the system, that is, all but a subset that has zero weight for the invariant measure. Our first task, in Section 1.1, will be to explain what we mean by 'dynamical system' and 'invariant measure'.

The roots of the theory date back to the first half of the 19th century. By 1838, the French mathematician Joseph Liouville observed that every energy-preserving system in Classical (Newtonian) Mechanics admits a natural invariant volume measure in the space of configurations. Just a bit later, in 1845, the great German mathematician Carl Friedrich Gauss pointed out that the transformation

$$(0,1] \to \mathbb{R}, \quad x \mapsto \text{fractional part of } \frac{1}{x},$$

which has an important role in Number Theory, admits an invariant measure equivalent to the Lebesgue measure (in the sense that the two have the same zero measure sets). These are two of the examples of applications of Ergodic Theory that we discuss in Section 1.3. Many others are introduced throughout this book.

The first important result was found by the great French mathematician Henri Poincaré by the end of the 19th century. Poincaré was particularly interested in the motion of celestial bodies, such as planets and comets, which is described by certain differential equations originating from Newton's Law of Universal Gravitation. Starting from Liouville's observation, Poincaré realized that for almost every initial state of the system, that is, almost every value of the initial position and velocity, the solution of the differential equation comes back arbitrarily close to that initial state, unless it goes to infinity. Even more, this *recurrence* property is not specific to (Celestial) Mechanics: it is shared by any dynamical system that admits a finite invariant measure. That is the theme of Section 1.2.

The same theme reappears in Section 1.5, in a more elaborate context: there, we deal with any finite number of dynamical systems commuting with each other, and we seek *simultaneous* returns of the orbits of all those systems to the neighborhood of the initial state. This kind of result has important applications in Combinatorics and Number Theory, as we will see.

The recurrence phenomenon is also behind the constructions that we present in Section 1.4. The basic idea is to fix some positive measure subset of the domain and to consider the first return to that subset. This first-return transformation is often easier to analyze, and it may be used to shed much light on the behavior of the original transformation.

1.1 Invariant measures

Let $(M, \mathcal{B}, \mu)$ be a measure space and $f : M \to M$ be a measurable transformation. We say that the measure μ is *invariant* under f if

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{for every measurable set } E \subset M. \tag{1.1.1}$$

We also say that μ is *f-invariant*, or that f *preserves* μ, to mean just the same. Notice that the definition (1.1.1) makes sense, since the pre-image of a measurable set under a measurable transformation is still a measurable set. Heuristically, the definition means that the probability that a point picked "at random" is in a given subset is equal to the probability that its image is in that subset.

It is possible, and convenient, to extend this definition to other types of dynamical systems, beyond transformations. We are especially interested in *flows*, that is, families of transformations $f^t : M \to M$, with $t \in \mathbb{R}$, satisfying the following conditions:

$$f^0 = \text{id} \quad \text{and} \quad f^{s+t} = f^s \circ f^t \text{ for every } s, t \in \mathbb{R}. \tag{1.1.2}$$

In particular, each transformation f^t is invertible and the inverse is f^{-t}. Flows arise naturally in connection with differential equations of the form

$$\frac{d\gamma}{dt}(t) = X(\gamma(t))$$

in the following way: under suitable conditions on the vector field X, for each point x in the domain M there exists exactly one solution $t \mapsto \gamma_x(t)$ of the differential equation with $\gamma_x(0) = x$; then $f^t(x) = \gamma_x(t)$ defines a flow in M.

We say that a measure μ is *invariant* under a flow $(f^t)_t$ if it is invariant under each one of the transformations f^t, that is, if

$$\mu(E) = \mu(f^{-t}(E)) \quad \text{for every measurable set } E \subset M \text{ and } t \in \mathbb{R}. \tag{1.1.3}$$

Proposition 1.1.1. *Let $f : M \to M$ be a measurable transformation and μ be a measure on M. Then f preserves μ if and only if*

$$\int \phi \, d\mu = \int \phi \circ f \, d\mu \tag{1.1.4}$$

for every μ-integrable function $\phi : M \to \mathbb{R}$.

Proof. Suppose that the measure μ is invariant under f. We are going to show that the relation (1.1.4) is valid for increasingly broader classes of functions. Let $\mathcal{X}_B$ denote the characteristic function of any measurable subset B. Then

$$\mu(B) = \int \mathcal{X}_B \, d\mu \quad \text{and} \quad \mu(f^{-1}(B)) = \int \mathcal{X}_{f^{-1}(B)} \, d\mu = \int (\mathcal{X}_B \circ f) \, d\mu.$$

Thus, the hypothesis $\mu(B) = \mu(f^{-1}(B))$ means that (1.1.4) is valid for characteristic functions. Then, by linearity of the integral, (1.1.4) is valid for all simple functions. Next, given any integrable $\phi : M \to \mathbb{R}$, consider a sequence $(s_n)_n$ of simple functions, converging to ϕ and such that $|s_n| \le |\phi|$ for every n. That such a sequence exists is guaranteed by Proposition A.1.33. Then, using the dominated convergence theorem (Theorem A.2.11) twice:

$$\int \phi \, d\mu = \lim_n \int s_n \, d\mu = \lim_n \int (s_n \circ f) \, d\mu = \int (\phi \circ f) \, d\mu.$$

This shows that (1.1.4) holds for every integrable function if μ is invariant. The converse is also contained in the arguments we just presented. □

1.1.1 Exercises

1.1.1. Let $f : M \to M$ be a measurable transformation. Show that a Dirac measure δ_p is invariant under f if and only if p is a fixed point of f. More generally, a probability measure $\delta_{p,k} = k^{-1}\big(\delta_p + \delta_{f(p)} + \cdots + \delta_{f^{k-1}(p)}\big)$ is invariant under f if and only if $f^k(p) = p$.

1.1.2. Prove the following version of Proposition 1.1.1. Let M be a metric space, $f : M \to M$ be a measurable transformation and μ be a measure on M. Show that f preserves μ if and only if

$$\int \phi \, d\mu = \int \phi \circ f \, d\mu$$

for every bounded continuous function $\phi : M \to \mathbb{R}$.

1.1.3. Prove that if $f : M \to M$ preserves a measure μ then, given any $k \ge 2$, the iterate f^k also preserves μ. Is the converse true?

1.1.4. Suppose that $f : M \to M$ preserves a probability measure μ. Let $B \subset M$ be a measurable set satisfying any one of the following conditions:

1. $\mu(B \setminus f^{-1}(B)) = 0$;
2. $\mu(f^{-1}(B) \setminus B) = 0$;
3. $\mu(B \Delta f^{-1}(B)) = 0$;
4. $f(B) \subset B$.

Show that there exists $C \subset M$ such that $f^{-1}(C) = C$ and $\mu(B \Delta C) = 0$.

1.1.5. Let $f : U \to U$ be a C^1 diffeomorphism on an open set $U \subset \mathbb{R}^d$. Show that the Lebesgue measure m is invariant under f if and only if $|\det Df| \equiv 1$.

1.2 Poincaré recurrence theorem

We are going to study two versions of Poincaré's theorem. The first one (Section 1.2.1) is formulated in the context of (finite) measure spaces. The theorem of Kač, that we state and prove in Section 1.2.2, provides a quantitative complement to that statement. The second version of the recurrence theorem (Section 1.2.3) assumes that the ambient is a topological space with certain additional properties. We will also prove a third version of the recurrence theorem, due to Birkhoff, whose statement is purely topological.

1.2.1 Measurable version

Our first result asserts that, given any *finite* invariant measure, almost every point in any positive measure set E returns to E an infinite number of times:

Theorem 1.2.1 (Poincaré recurrence)**.** *Let $f : M \to M$ be a measurable transformation and μ be a finite measure invariant under f. Let $E \subset M$ be any measurable set with $\mu(E) > 0$. Then, for μ-almost every point $x \in E$ there exist infinitely many values of n for which $f^n(x)$ is also in E.*

Proof. Denote by E_0 the set of points $x \in E$ that never return to E. As a first step, let us prove that E_0 has zero measure. To this end, let us observe that the pre-images $f^{-n}(E_0)$ are pairwise disjoint. Indeed, suppose there exist $m > n \geq 1$ such that $f^{-m}(E_0)$ intersects $f^{-n}(E_0)$. Let x be a point in the intersection and $y = f^n(x)$. Then $y \in E_0$ and $f^{m-n}(y) = f^m(x) \in E_0$. Since $E_0 \subset E$, this means that y returns to E at least once, which contradicts the definition of E_0. This contradiction proves that the pre-images are pairwise disjoint, as claimed.

Since μ is invariant, we also have that $\mu(f^{-n}(E_0)) = \mu(E_0)$ for all $n \geq 1$. It follows that

$$\mu\Big(\bigcup_{n=1}^{\infty} f^{-n}(E_0)\Big) = \sum_{n=1}^{\infty} \mu(f^{-n}(E_0)) = \sum_{n=1}^{\infty} \mu(E_0).$$

The expression on the left-hand side is finite, since the measure μ is assumed to be finite. On the right-hand side we have a sum of infinitely many terms that are all equal. The only way such a sum can be finite is if the terms vanish. So, $\mu(E_0) = 0$ as claimed.

Now let us denote by F the set of points $x \in E$ that return to E a finite number of times. It is clear from the definition that every point $x \in F$ has some iterate $f^k(x)$ in E_0. In other words,

$$F \subset \bigcup_{k=0}^{\infty} f^{-k}(E_0).$$

Since $\mu(E_0) = 0$ and μ is invariant, it follows that

$$\mu(F) \leq \mu\Big(\bigcup_{k=0}^{\infty} f^{-k}(E_0)\Big) \leq \sum_{k=0}^{\infty} \mu\big(f^{-k}(E_0)\big) = \sum_{k=0}^{\infty} \mu(E_0) = 0.$$

Thus, $\mu(F) = 0$ as we wanted to prove. □

Theorem 1.2.1 implies an analogous result for continuous time systems: *if μ is a finite invariant measure of a flow $(f^t)_t$ then for every measurable set $E \subset M$ with positive measure and for μ-almost every $x \in E$ there exist times $t_j \to +\infty$ such that $f^{t_j}(x) \in E$.* Indeed, if μ is invariant under the flow then, in particular, it is invariant under the so-called *time-1 map* f^1. So, the statement we just made follows immediately from Theorem 1.2.1 applied to f^1 (the times t_j one finds in this way are integers). Similar observations apply to the other versions of the recurrence theorem that we present in the sequel.

On the other hand, the theorem in the next section is specific to discrete time systems.

1.2.2 Kač theorem

Let $f : M \to M$ be a measurable transformation and μ be a finite measure invariant under f. Let $E \subset M$ be any measurable set with $\mu(E) > 0$. Consider the *first-return time* function $\rho_E : E \to \mathbb{N} \cup \{\infty\}$, defined by

$$\rho_E(x) = \min\{n \geq 1 : f^n(x) \in E\} \tag{1.2.1}$$

if the set on the right-hand side is non-empty and $\rho_E(x) = \infty$ if, on the contrary, x has no iterate in E. According to Theorem 1.2.1, the second alternative occurs only on a set with zero measure.

The next result shows that this function is integrable and even provides the value of the integral. For the statement we need the following notation:

$$E_0 = \{x \in E : f^n(x) \notin E \text{ for every } n \geq 1\} \quad \text{and}$$
$$E_0^* = \{x \in M : f^n(x) \notin E \text{ for every } n \geq 0\}.$$

In other words, E_0 is the set of points in E that never *return* to E and E_0^* is the set of points in M that never *enter* E. We have seen in Theorem 1.2.1 that $\mu(E_0) = 0$.

Theorem 1.2.2 (Kač)**.** *Let $f : M \to M$ be a measurable transformation, μ be a finite invariant measure and $E \subset M$ be a positive measure set. Then the function ρ_E is integrable and*

$$\int_E \rho_E \, d\mu = \mu(M) - \mu(E_0^*).$$

Proof. For each $n \geq 1$, define

$$E_n = \{x \in E : f(x) \notin E, \dots, f^{n-1}(x) \notin E, \text{ but } f^n(x) \in E\} \quad \text{and}$$
$$E_n^* = \{x \in M : x \notin E, f(x) \notin E, \dots, f^{n-1}(x) \notin E, \text{ but } f^n(x) \in E\}.$$

That is, E_n is the set of points of E that return to E for the first time exactly at time n,

$$E_n = \{x \in E : \rho_E(x) = n\},$$

and E_n^* is the set points that are not in E and enter E for the first time exactly at time n. It is clear that these sets are measurable and, hence, ρ_E is a measurable function. Moreover, the sets E_n, E_n^*, $n \geq 0$ constitute a *partition* of the ambient space: they are pairwise disjoint and their union is the whole of M. So,

$$\mu(M) = \sum_{n=0}^{\infty} \big(\mu(E_n) + \mu(E_n^*)\big) = \mu(E_0^*) + \sum_{n=1}^{\infty} \big(\mu(E_n) + \mu(E_n^*)\big). \tag{1.2.2}$$

Now observe that

$$f^{-1}(E_n^*) = E_{n+1}^* \cup E_{n+1} \quad \text{for every } n. \tag{1.2.3}$$

Indeed, $f(y) \in E_n^*$ means that the first iterate of $f(y)$ that belongs to E is $f^n(f(y)) = f^{n+1}(y)$ and that occurs if and only if $y \in E_{n+1}^*$ or else $y \in E_{n+1}$. This proves the equality (1.2.3). So, given that μ is invariant,

$$\mu(E_n^*) = \mu(f^{-1}(E_n^*)) = \mu(E_{n+1}^*) + \mu(E_{n+1}) \quad \text{for every } n.$$

Applying this relation successively, we find that

$$\mu(E_n^*) = \mu(E_m^*) + \sum_{i=n+1}^{m} \mu(E_i) \quad \text{for every } m > n. \tag{1.2.4}$$

The relation (1.2.2) implies that $\mu(E_m^*) \to 0$ when $m \to \infty$. So, taking the limit as $m \to \infty$ in the equality (1.2.4), we find that

$$\mu(E_n^*) = \sum_{i=n+1}^{\infty} \mu(E_i). \tag{1.2.5}$$

To complete the proof, replace (1.2.5) in the equality (1.2.2). In this way we find that

$$\mu(M) - \mu(E_0^*) = \sum_{n=1}^{\infty} \Big(\sum_{i=n}^{\infty} \mu(E_i)\Big) = \sum_{n=1}^{\infty} n\mu(E_n) = \int_E \rho_E \, d\mu,$$

as we wanted to prove. □

In some cases, for example when the system (f, μ) is *ergodic* (this property will be defined and studied later, starting from Chapter 4), the set E_0^* has zero measure. Then the conclusion of the Kač theorem means that

$$\frac{1}{\mu(E)} \int_E \rho_E \, d\mu = \frac{\mu(M)}{\mu(E)} \tag{1.2.6}$$

for every measurable set E with positive measure. The left-hand side of this expression is the *mean return time* to E. So, (1.2.6) asserts that *the mean return time is inversely proportional to the measure of E.*

Remark 1.2.3. By definition, $E_n^* = f^{-n}(E) \setminus \cup_{k=0}^{n-1} f^{-k}(E)$. So, the fact that the sum (1.2.2) is finite implies that the measure of E_n^* converges to zero when $n \to \infty$. This fact will be useful later.

1.2.3 Topological version

Now let us suppose that M is a topological space, endowed with its Borel σ-algebra $\mathcal{B}$. A point $x \in M$ is *recurrent* for a transformation $f : M \to M$ if there exists a sequence $n_j \to \infty$ of natural numbers such that $f^{n_j}(x) \to x$. Analogously, we say that $x \in M$ is recurrent for a flow $(f^t)_t$ if there exists a sequence $t_j \to +\infty$ of real numbers such that $f^{t_j}(x) \to x$ when $j \to \infty$.

In the next theorem we assume that the topological space M admits a countable basis of open sets, that is, there exists a countable family $\{U_k : k \in \mathbb{N}\}$ of open sets such that every open subset of M may be written as a union of elements U_k of this family. This condition holds in most interesting examples.

Theorem 1.2.4 (Poincaré recurrence)**.** *Suppose that M admits a countable basis of open sets. Let $f : M \to M$ be a measurable transformation and μ be a finite measure on M invariant under f. Then, μ-almost every $x \in M$ is recurrent for f.*

Proof. For each k, denote by $\tilde{U}_k$ the set of points $x \in U_k$ that never return to U_k. According to Theorem 1.2.1, every $\tilde{U}_k$ has zero measure. Consequently, the countable union

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} \tilde{U}_k$$

also has zero measure. Hence, to prove the theorem it suffices to check that every point x that is not in $\tilde{U}$ is recurrent. That is easy, as we are going to see. Consider $x \in M \setminus \tilde{U}$ and let U be any neighborhood of x. By definition, there exists some element U_k of the basis of open sets such that $x \in U_k$ and $U_k \subset U$. Since x is not in $\tilde{U}$, we also have that $x \notin \tilde{U}_k$. In other words, there exists $n \geq 1$ such that $f^n(x)$ is in U_k. In particular, $f^n(x)$ is also in U. Since the neighborhood U is arbitrary, this proves that x is a recurrent point. □

Let us point out that the conclusions of Theorems 1.2.1 and 1.2.4 are false, in general, if the measure μ is not finite:

Example 1.2.5. Let $f : \mathbb{R} \to \mathbb{R}$ be the translation by 1, that is, the transformation defined by $f(x) = x + 1$ for every $x \in \mathbb{R}$. It is easy to check that f preserves the Lebesgue measure on $\mathbb{R}$ (which is infinite). On the other hand, no point $x \in \mathbb{R}$ is recurrent for f. According to the recurrence theorem, this last observation implies that f can not admit any finite invariant measure.

However, it is possible to extend these statements for certain cases of infinite measures: see Exercise 1.2.2.

To conclude, we present a purely topological version of Theorem 1.2.4, called the Birkhoff recurrence theorem, that makes no reference at all to invariant measures:

Theorem 1.2.6 (Birkhoff recurrence)**.** *If $f : M \to M$ is a continuous transformation on a compact metric space M then there exists some point $x \in M$ that is recurrent for f.*

Proof. Consider the family $\mathcal{I}$ of all non-empty closed sets $X \subset M$ that are invariant under f, in the sense that $f(X) \subset X$. This family is non-empty, since $M \in \mathcal{I}$. We claim that an element $X \in \mathcal{I}$ is minimal for the inclusion relation if and only if the orbit of every $x \in X$ is dense in X. Indeed, it is clear that if X is a closed invariant subset then X contains the closure of the orbit of each one of its elements. Hence, in order to be minimal, X must coincide with every one of these closures. Conversely, for the same reason, if X coincides with the orbit closure of each one of its points then it has no proper subset that is closed and invariant. That is, X is minimal. This proves our claim. In particular, every point x in a minimal set is recurrent. Therefore, to prove the theorem it suffices to prove that there exists some minimal set.

We claim that every totally ordered set $\{X_\alpha\} \subset \mathcal{I}$ admits a lower bound. Indeed, consider $X = \cap_\alpha X_\alpha$. Observe that X is non-empty, since the X_α are compact and they form a totally ordered family. It is clear that X is closed and invariant under f and it is equally clear that X is a lower bound for the set $\{X_\alpha\}$. That proves our claim. Now it follows from Zorn's lemma that $\mathcal{I}$ does contain minimal elements. □

Theorem 1.2.6 can also be deduced from Theorem 1.2.4 together with the fact, which we will prove later (in Chapter 2), that every continuous transformation on a compact metric space admits some invariant probability measure.

1.2.4 Exercises

1.2.1. Show that the following statement is equivalent to Theorem 1.2.1, meaning that each one of them can be obtained from the other. Let $f : M \to M$ be a measurable transformation and μ be a finite invariant measure. Let $E \subset M$ be any measurable set with $\mu(E) > 0$. Then there exists $N \geq 1$ and a positive measure set $D \subset E$ such that $f^N(x) \in E$ for every $x \in D$.

1.2.2. Let $f : M \to M$ be an invertible transformation and suppose that μ is an invariant measure, not necessarily finite. Let $B \subset M$ be a set with finite measure. Prove that, given any measurable set $E \subset M$ with positive measure, μ-almost every point $x \in E$ either returns to E an infinite number of times or has only a finite number of iterates in B.

1.2.3. Let $f : M \to M$ be an invertible transformation and suppose that μ is a *σ-finite* invariant measure: there exists an increasing sequence of measurable subsets M_k with $\mu(M_k) < \infty$ for every k and $\cup_k M_k = M$. We say that a point *x goes to infinity* if, for every k, there exists only a finite number of iterates of x that are in M_k. Show that, given any $E \subset M$ with positive measure, μ-almost every point $x \in E$ returns to E an infinite number of times or else goes to infinity.

1.2.4. Let $f : M \to M$ be a measurable transformation, not necessarily invertible, μ be an invariant probability measure and $D \subset M$ be a set with positive measure. Prove that almost every point of D spends a *positive fraction of time* in D:

$$\limsup_n \frac{1}{n}\#\{0 \le j \le n-1 : f^j(x) \in D\} > 0$$

for μ-almost every $x \in D$. [Note: One may replace lim sup by lim inf in the statement, but the proof of that fact will have to wait until Chapter 3.]

1.2.5. Let $f : M \to M$ be a measurable transformation preserving a finite measure μ. Given any measurable set $A \subset M$ with $\mu(A) > 0$, let $n_1 < n_2 < \cdots$ be the sequence of values of n such that $\mu(f^{-n}(A) \cap A) > 0$. The goal of this exercise is to prove that $V_A = \{n_1, n_2, \dots\}$ is a *syndetic*, that is, that there exists $C > 0$ such that $n_{i+1} - n_i \le C$ for every i.

1. Show that for any increasing sequence $k_1 < k_2 < \cdots$ there exist $j > i \ge 1$ such that $\mu(A \cap f^{-(k_j - k_i)}(A)) > 0$.

2. Given any infinite sequence $\ell = (l_j)_j$ of natural numbers, denote by $S(\ell)$ the set of all finite sums of consecutive elements of ℓ. Show that V_A intersects $S(\ell)$ for every ℓ.

3. Deduce that the set V_A is syndetic.

[Note: Exercise 3.1.2 provides a different proof of this fact.]

1.2.6. Show that if $f : [0,1] \to [0,1]$ is a measurable transformation preserving the Lebesgue measure m then m-almost every point $x \in [0,1]$ satisfies

$$\liminf_n n|f^n(x) - x| \le 1.$$

[Note: Boshernitzan [Bos93] proved a much more general result, namely that $\liminf_n n^{1/d} d(f^n(x), x) < \infty$ for μ-almost every point and every probability measure μ invariant under $f : M \to M$, assuming M is a separable metric whose d-dimensional Hausdorff measure is σ-finite.]

1.2.7. Define $f : [0,1] \to [0,1]$ by $f(x) = (x+\omega) - [x+\omega]$, where ω represents the *golden ratio* $(1+\sqrt{5})/2$. Given $x \in [0,1]$, check that $n|f^n(x) - x| = n^2|\omega - q_n|$ for every n, where $(q_n)_n \to \omega$ is the sequence of rational numbers given by $q_n = [x + n\omega]/n$. Using that the roots of the polynomial $R(z) = z^2 - z - 1$ are precisely ω and $\omega - \sqrt{5}$, prove that $\liminf_n n^2|\omega - q_n| \ge 1/\sqrt{5}$. [Note: This shows that the constant 1 in Exercise 1.2.6 cannot be replaced by any constant smaller than $1/\sqrt{5}$. It is not known whether 1 is the smallest constant such that the statement holds for *every* transformation on the interval.]

1.3 Examples

Next, we describe some simple examples of invariant measures for transformations and flows that help us interpret the significance of the Poincaré recurrence theorem and also lead to some interesting conclusions.

1.3.1 Decimal expansion

Our first example is the transformation defined on the interval $[0, 1]$ in the following way:

$$f : [0, 1] \to [0, 1], \quad f(x) = 10x - [10x].$$

Here and in what follows, we use $[y]$ as the *integer part* of a real number y, that is, the largest integer smaller than or equal to y. So, f is the map sending each $x \in [0, 1]$ to the *fractional part* of $10x$. Figure 1.1 represents the graph of f.

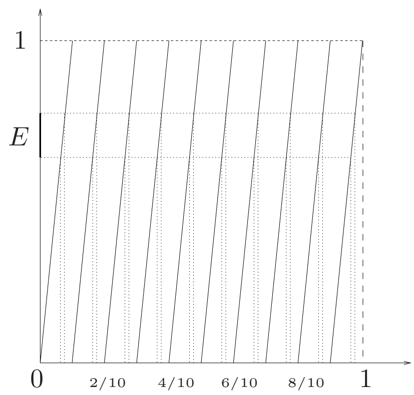

Figure 1.1: Fractional part of $10x$

We claim that the Lebesgue measure μ on the interval is invariant under the transformation f, that is, it satisfies

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{for every measurable set } E \subset M. \tag{1.3.1}$$

This can be checked as follows. Let us begin by supposing that E is an interval. Then, as illustrated in Figure 1.1, its pre-image $f^{-1}(E)$ consists of ten intervals, each of which is ten times shorter than E. Hence, the Lebesgue measure of $f^{-1}(E)$ is equal to the Lebesgue measure of E. This proves that (1.3.1) does hold in the case of intervals. As a consequence, it also holds when E is a finite union of intervals. Now, the family of all finite unions of intervals is an algebra that generates the Borel σ-algebra of $[0, 1]$. Hence, to conclude the proof it is enough to use the following general fact:

Lemma 1.3.1. *Let $f : M \to M$ be a measurable transformation and μ be a finite measure on M. Suppose that there exists some algebra $\mathcal{A}$ of measurable subsets of M such that $\mathcal{A}$ generates the σ-algebra $\mathcal{B}$ of M and $\mu(E) = \mu(f^{-1}(E))$ for every $E \in \mathcal{A}$. Then the latter remains true for every set $E \in \mathcal{B}$, that is, the measure μ is invariant under f.*

Proof. We start by proving that $\mathcal{C} = \{E \in \mathcal{B} : \mu(E) = \mu(f^{-1}(E))\}$ is a monotone class. Let $E_1 \subset E_2 \subset \dots$ be any increasing sequence of elements of $\mathcal{C}$ and let $E = \cup_{i=1}^{\infty} E_i$. By Theorem A.1.14 (see Exercise A.1.9),

$$\mu(E) = \lim_i \mu(E_i) \quad \text{and} \quad \mu(f^{-1}(E)) = \lim_i \mu(f^{-1}(E_i)).$$

So, using the fact that $E_i \in \mathcal{C}$,

$$\mu(E) = \lim_i \mu(E_i) = \lim_i \mu(f^{-1}(E_i)) = \mu(f^{-1}(E)).$$

Hence, $E \in \mathcal{C}$. In precisely the same way, one gets that the intersection of any decreasing sequence of elements of $\mathcal{C}$ is in $\mathcal{C}$. This proves that $\mathcal{C}$ is indeed a monotone class.

Now it is easy to deduce the conclusion of the lemma. Indeed, since $\mathcal{C}$ is assumed to contain $\mathcal{A}$, we may use the monotone class theorem (Theorem A.1.18), to conclude that $\mathcal{C}$ contains the σ-algebra $\mathcal{B}$ generated by $\mathcal{A}$. That is precisely what we wanted to prove. □

Now we explain how one may use the fact that the Lebesgue measure is invariant under f, together with the Poincaré recurrence theorem, to reach some interesting conclusions. The transformation f is directly related to the usual decimal expansion of a real number: if x is given by

$$x = 0.a_0 a_1 a_2 a_3 \cdots$$

with $a_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $a_i \neq 9$ for infinitely many values of i, then its image under f is given by

$$f(x) = 0.a_1 a_2 a_3 \cdots .$$

Thus, more generally, the n-th iterate of f can be expressed in the following way, for every $n \geq 1$:

$$f^n(x) = 0.a_n a_{n+1} a_{n+2} \cdots \qquad (1.3.2)$$

Let E be the subset of points $x \in [0,1]$ whose decimal expansion starts with the digit 7, that is, such that $a_0 = 7$. According to Theorem 1.2.1, almost every element in E has infinitely many iterates that are also in E. By the expression (1.3.2), this means that there are infinitely many values of n such that $a_n = 7$. So, we have shown that *almost every number x whose decimal expansion starts with* 7 *has infinitely many digits equal to* 7.

Of course, instead of 7 we may consider any other digit. Even more, there is a similar result (see Exercise 1.3.2) when, instead of a single digit, one considers a block of $k \geq 1$ consecutive digits. Later on, in Chapter 3, we will prove a much stronger fact: for almost every number $x \in [0,1]$, every digit occurs with frequency $1/10$ (more generally, every block of $k \geq 1$ digits occurs with frequency $1/10^k$) in the decimal expansion of x.

1.3.2 Gauss map

The system we present in this section is related to another important algorithm in Number Theory, the *continued fraction expansion*, which plays a central role in the problem of finding the best rational approximation to any real number. Let us start with a brief presentation of this algorithm.

Given any number $x_0 \in (0,1)$, let

$$a_1 = \left[\frac{1}{x_0}\right] \qquad \text{and} \qquad x_1 = \frac{1}{x_0} - a_1.$$

Note that a_1 is a natural number, $x_1 \in [0,1)$ and

$$x_0 = \frac{1}{a_1 + x_1}.$$

Supposing that x_1 is different from zero, we may repeat this procedure, defining

$$a_2 = \left[\frac{1}{x_1}\right] \qquad \text{and} \qquad x_2 = \frac{1}{x_1} - a_2.$$

Then

$$x_1 = \frac{1}{a_1 + x_2} \quad \text{and so} \quad x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + x_2}}.$$

Now we may proceed by induction: for each $n \geq 1$ such that $x_{n-1} \in (0,1)$, define

$$a_n = \left[\frac{1}{x_{n-1}}\right] \qquad \text{and} \qquad x_n = \frac{1}{x_{n-1}} - a_n = G(x_{n-1}),$$

and observe that

$$x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n + x_n}}}}. \tag{1.3.3}$$

It can be shown that the sequence

$$z_n = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}} \tag{1.3.4}$$

converges to x_0 when $n \to \infty$. This is usually expressed through the expression

$$x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n + \cfrac{1}{\cdots}}}}}, \tag{1.3.5}$$

which is called *continued fraction expansion* of x_0.

Note that the sequence $(z_n)_n$ defined by the relation (1.3.4) consists of rational numbers. Indeed, one can show that these are the *best rational approximations* of the number x_0, in the sense that each z_n is closer to x_0 than any other rational number whose denominator is smaller than or equal to the denominator of z_n (written in irreducible form). Observe also that to obtain (1.3.5) we had to assume that $x_n \in (0,1)$ for every $n \in \mathbb{N}$. If in the course of the process one encounters some $x_n = 0$, then the algorithm halts and we consider (1.3.3) to be the continued fraction expansion of x_0. It is clear that this can happen only if x_0 itself is a rational number.

This continued fraction algorithm is intimately related to a certain dynamical system on the interval $[0,1]$ that we describe in the following. The *Gauss map* $G : [0,1] \to [0,1]$ is defined by

$$G(x) = \frac{1}{x} - \left[\frac{1}{x}\right] = \text{fractional part of } 1/x,$$

if $x \in (0,1]$ and $G(0) = 0$. The graph of G can be easily sketched (see Figure 1.2), starting from the following observation: for every x in each interval $I_k = (1/(k+1), 1/k]$, the integer part of $1/x$ is equal to k and so $G(x) = 1/x - k$.

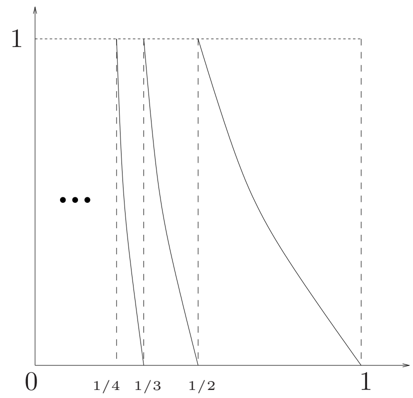

Figure 1.2: Gauss map

The continued fraction expansion of any number $x_0 \in (0,1)$ can be obtained from the Gauss map, in the following way: for each $n \geq 1$, the natural number a_n is determined by

$$G^{n-1}(x_0) \in I_{a_n},$$

and the real number x_n is simply the n-th iterate $G^n(x_0)$ of the point x_0. This process halts whenever we encounter some $x_n = 0$; as we explained previously, this can only happen if x_0 is a rational number (see Exercise 1.3.4). In particular, there exists a full Lebesgue measure subset of $(0,1)$ such that all the iterates of G are defined for all the points in that subset.

A remarkable fact that makes this transformation interesting from the point of view of Ergodic Theory is that G admits an invariant probability measure that, in addition, is equivalent to the Lebesgue measure on the interval. Indeed, consider the measure defined by

$$\mu(E) = \int_E \frac{c}{1+x}\, dx \quad \text{for every measurable set } E \subset [0,1], \tag{1.3.6}$$

where c is a positive constant. The integral is well defined, since the function in the integral is continuous on the interval $[0,1]$. Moreover, this function takes values inside the interval $[c/2, c]$ and that implies

$$\frac{c}{2}\, m(E) \le \mu(E) \le c\, m(E) \quad \text{for every measurable set } E \subset [0,1]. \tag{1.3.7}$$

In particular, μ is indeed equivalent to the Lebesgue measure m.

Proposition 1.3.2. *The measure μ is invariant under G. Moreover, if we choose $c = 1/\log 2$ then μ is a probability measure.*

Proof. We are going to use the following lemma:

Lemma 1.3.3. *Let $f : [0,1] \to [0,1]$ be a transformation such that there exist pairwise disjoint open intervals $I_1, I_2, \ldots$ satisfying*

1. *the union $\cup_k I_k$ has full Lebesgue measure in $[0,1]$ and*
2. *the restriction $f_k = f \mid I_k$ to each I_k is a diffeomorphism onto $(0,1)$.*

Let $\rho : [0,1] \to [0,\infty)$ be an integrable function (relative to the Lebesgue measure) satisfying

$$\rho(y) = \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|f'(x)|} \tag{1.3.8}$$

for almost every $y \in [0,1]$. Then the measure $\mu = \rho dx$ is invariant under f.

Proof. Let $\phi = \chi_E$ be the characteristic function of an arbitrary measurable set $E \subset [0,1]$. Changing variables in the integral,

$$\int_{I_k} \phi(f(x))\rho(x)\, dx = \int_0^1 \phi(y)\rho(f_k^{-1}(y))|(f_k^{-1})'(y)|\, dy.$$

Note that $(f_k^{-1})'(y) = 1/f'(f_k^{-1}(y))$. So, the previous relation implies that

$$\begin{aligned} \int_0^1 \phi(f(x))\rho(x)\, dx &= \sum_{k=1}^{\infty} \int_{I_k} \phi(f(x))\rho(x)\, dx \\ &= \sum_{k=1}^{\infty} \int_0^1 \phi(y) \frac{\rho(f_k^{-1}(y))}{|f'(f_k^{-1}(y))|}\, dy. \end{aligned} \tag{1.3.9}$$

Using the monotone convergence theorem (Theorem A.2.9) and the hypothesis (1.3.8), we see that the last expression in (1.3.9) is equal to

$$\int_0^1 \phi(y) \sum_{k=1}^{\infty} \frac{\rho(f_k^{-1}(y))}{|f'(f_k^{-1}(y))|}\, dy = \int_0^1 \phi(y)\rho(y)\, dy.$$

In this way we find that $\int_0^1 \phi(f(x))\rho(x)\,dx = \int_0^1 \phi(y)\rho(y)\,dy$. Since $\mu = \rho dx$ and $\phi = \mathcal{X}_E$, this means that $\mu(f^{-1}(E)) = \mu(E)$ for every measurable set $E \subset [0,1]$. In other words, μ is invariant under f. □

To conclude the proof of Proposition 1.3.2 we must show that the condition (1.3.8) holds for $\rho(x) = c/(1+x)$ and $f = G$. Let G_k denote the restriction of G to the interval $I_k = (1/(k+1), 1/k)$, for $k \geq 1$. Note that $G_k^{-1}(y) = 1/(y+k)$ for every k. Note also that $G'(x) = (1/x)' = -1/x^2$ for every $x \neq 0$. Therefore,

$$\sum_{k=1}^{\infty} \frac{\rho(G_k^{-1}(y))}{|G'(G_k^{-1}(y))|} = \sum_{k=1}^{\infty} \frac{c(y+k)}{y+k+1}\Big(\frac{1}{y+k}\Big)^2 = \sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)}. \quad (1.3.10)$$

Observing that

$$\frac{1}{(y+k)(y+k+1)} = \frac{1}{y+k} - \frac{1}{y+k+1}\,,$$

we see that the last sum in (1.3.10) has a telescopic structure: except for the first one, all the terms occur twice, with opposite signs, and so they cancel out. This means that the sum is equal to the first term:

$$\sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \frac{c}{y+1} = \rho(y).$$

This proves that the equality (1.3.8) is indeed satisfied and, hence, we may use Lemma 1.3.1 to conclude that μ is invariant under f.

Finally, observing that $c\log(1+x)$ is a primitive of the function $\rho(x)$, we find that

$$\mu([0,1]) = \int_0^1 \frac{c}{1+x}\, dx = c\log 2.$$

So, picking $c = 1/\log 2$ ensures that μ is a probability measure. □

This proposition allows us to use ideas from Ergodic Theory, applied to the Gauss map, to obtain interesting conclusions in Number Theory. For example (see Exercise 1.3.3), the natural number 7 occurs infinitely many times in the continued fraction expansion of almost every number $x_0 \in (1/8, 1/7)$, that is, one has $a_n = 7$ for infinitely many values of $n \in \mathbb{N}$. Later on, in Chapter 3, we will prove a much more precise statement, that contains the following conclusion: for almost every $x_0 \in (0,1)$ the number 7 occurs with frequency

$$\frac{1}{\log 2} \log \frac{64}{63}$$

in the continued fraction expansion of x_0. Try to guess right away where this number comes from!

1.3.3 Circle rotations

Let us consider on the real line $\mathbb{R}$ the equivalence relation $\sim$ that identifies any numbers whose difference is an integer number:

$$x \sim y \quad \Leftrightarrow \quad x - y \in \mathbb{Z}.$$

We represent by $[x] \in \mathbb{R}/\mathbb{Z}$ the equivalence class of each $x \in \mathbb{R}$ and denote by $\mathbb{R}/\mathbb{Z}$ the space of all equivalence classes. This space is called the *circle* and is also denoted by S^1. The reason for this terminology is that $\mathbb{R}/\mathbb{Z}$ can be identified in a natural way with the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ on the complex plane, by means of the map

$$\phi : \mathbb{R}/\mathbb{Z} \to \{z \in \mathbb{C} : |z| = 1\}, \quad [x] \mapsto e^{2\pi x i}. \tag{1.3.11}$$

Note that ϕ is well defined: since the function $x \mapsto e^{2\pi x i}$ is periodic of period 1, the expression $e^{2\pi x i}$ does not depend on the choice of a representative x for the class $[x]$. Moreover, ϕ is a bijection.

The circle $\mathbb{R}/\mathbb{Z}$ inherits from the real line $\mathbb{R}$ the structure of an abelian group, given by the operation

$$[x] + [y] = [x + y].$$

Observe that this is well defined: the equivalence class on the right-hand side does not depend on the choice of representatives x and y for the classes on the left-hand side. Given $\theta \in \mathbb{R}$, we call *rotation* of angle θ the transformation

$$R_\theta : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}, \quad [x] \mapsto [x + \theta] = [x] + [\theta].$$

Note that R_θ corresponds, via the identification (1.3.11), to the transformation $z \mapsto e^{2\pi\theta i} z$ on $\{z \in \mathbb{C} : |z| = 1\}$. The latter is just the restriction to the unit circle of the rotation of angle $2\pi\theta$ around the origin in the complex plane. It is clear from the definition that R_0 is the identity map and $R_\theta \circ R_\tau = R_{\theta+\tau}$ for every θ and τ. In particular, every R_θ is invertible and the inverse is $R_{-\theta}$.

We can also endow S^1 with a natural structure of a probability space, as follows. Let $\pi : \mathbb{R} \to S^1$ be the canonical projection, that assigns to each $x \in \mathbb{R}$ its equivalence class $[x]$. We say that a set $E \subset S^1$ is measurable if $\pi^{-1}(E)$ is a measurable subset of the real line. Next, let m be the Lebesgue measure on the real line. We define the *Lebesgue measure* μ on the circle to be given by

$$\mu(E) = m\big(\pi^{-1}(E) \cap [k, k+1)\big) \quad \text{for every } k \in \mathbb{Z}.$$

Note that the left-hand side of this equality does not depend on k, since, by definition, $\pi^{-1}(E) \cap [k, k+1) = \big(\pi^{-1}(E) \cap [0, 1)\big) + k$ and the measure m is invariant under translations.

It is clear from the definition that μ is a probability. Moreover, μ is invariant under every rotation R_θ (according to Exercise 1.3.8, it is the only probability measure with this property). This can be shown as follows. By definition,

$\pi^{-1}(R_\theta^{-1}(E)) = \pi^{-1}(E) - \theta$ for every measurable set $E \subset S^1$. Let k be the integer part of θ. Since m is invariant under all the translations,

$$\begin{aligned} m\big((\pi^{-1}(E) - \theta) \cap [0,1)\big) &= m\big(\pi^{-1}(E) \cap [\theta, \theta+1)\big) \\ &= m\big(\pi^{-1}(E) \cap [\theta, k+1)\big) + m\big(\pi^{-1}(E) \cap [k+1, \theta+1)\big). \end{aligned}$$

Note that $\pi^{-1}(E) \cap [k+1, \theta+1) = \big(\pi^{-1}(E) \cap [k, \theta)\big) + 1$. So, the expression on the right-hand side of the previous equality may be written as

$$m\big(\pi^{-1}(E) \cap [\theta, k+1)\big) + m\big(\pi^{-1}(E) \cap [k, \theta)\big) = m\big(\pi^{-1}(E) \cap [k, k+1)\big).$$

Combining these two relations we find that

$$\mu\big(R_\theta^{-1}(E)\big) = m\big(\pi^{-1}(R_\theta^{-1}(E) \cap [0,1))\big) = m\big(\pi^{-1}(E) \cap [k, k+1)\big) = \mu(E)$$

for every measurable set $E \subset S^1$.

The rotations $R_\theta : S^1 \to S^1$ exhibit two very different types of dynamical behavior, depending on the value of θ. If θ is rational, say $\theta = p/q$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$, then

$$R_\theta^q([x]) = [x + q\theta] = [x] \quad \text{for every } [x].$$

Consequently, in this case every point $x \in S^1$ is periodic with period q. In the opposite case we have:

Proposition 1.3.4. *If θ is irrational then $\mathcal{O}([x]) = \{R_\theta^n([x]) : n \in \mathbb{N}\}$ is a dense subset of the circle $\mathbb{R}/\mathbb{Z}$ for every $[x]$.*

Proof. We claim that the set $\mathcal{D} = \{m + n\theta : m \in \mathbb{Z}, n \in \mathbb{N}\}$ is dense in $\mathbb{R}$. Indeed, consider any number $r \in \mathbb{R}$. Given any $\varepsilon > 0$, we may choose $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that $|q\theta - p| < \varepsilon$. Note that the number $a = q\theta - p$ is necessarily different from zero, since θ is irrational. Let us suppose that a is positive (the case when a is negative is analogous). Subdividing the real line into intervals of length a, we see that there exists an integer l such that $0 \le r - la < a$. This implies that

$$|r - (lq\theta - lp)| = |r - la| < a < \varepsilon.$$

As $m = lq$ and $n = -lq$ are integers and ε is arbitrary, this proves that r is in the closure of the set $\mathcal{D}$, for every $r \in \mathbb{R}$.

Now, given $y \in \mathbb{R}$ and $\varepsilon > 0$, we may take $r = y - x$ and, using the previous paragraph, we may find $m, n \in \mathbb{Z}$ such that $|m + n\theta - (y - x)| < \varepsilon$. This is equivalent to saying that the distance from $[y]$ to the iterate $R_\theta^n([x])$ is less than ε. Since x, y and ε are arbitrary, this shows that every orbit $\mathcal{O}([x])$ is dense in S^1. □

In particular, it follows that *every* point on the circle is recurrent for R_θ (this is also true when θ is rational). The previous proposition also leads to some interesting conclusions in the study of the invariant measures of R_θ. Among other things, we will learn later (in Chapter 6) that if θ is irrational then the Lebesgue measure is the unique probability measure that is preserved by R_θ. Related to this, we will see that the orbits of R_θ are uniformly distributed subsets of S^1.

1.3.4 Rotations on tori

The notions we just presented can be generalized to arbitrary dimension, as we are going to explain. For each $d \geq 1$, consider the equivalence relation on $\mathbb{R}^d$ that identifies any two vectors whose difference is an integer vector:

$$(x_1, \ldots, x_d) \sim (y_1, \ldots, y_d) \quad \Leftrightarrow \quad (x_1 - y_1, \ldots, x_d - y_d) \in \mathbb{Z}^d.$$

We denote by $[x]$ or $[(x_1, \ldots, x_d)]$ the equivalence class of any $x = (x_1, \ldots, x_d)$. Then we call the *d-dimensional torus* or, simply, the *d-torus* the space

$$\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d = (\mathbb{R}/\mathbb{Z})^d$$

formed by those equivalence classes. Let m be the Lebesgue measure on $\mathbb{R}^d$. The operation

$$[(x_1, \ldots, x_d)] + [(y_1, \ldots, y_d)] = [(x_1 + y_1, \ldots, x_d + y_d)]$$

is well defined and turns $\mathbb{T}^d$ into an abelian group. Given $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$, we call

$$R_\theta : \mathbb{T}^d \to \mathbb{T}^d, \quad R_\theta([x]) = [x] + [\theta]$$

the *rotation* by θ (sometimes, R_θ is also called the *translation* by θ). The map

$$\phi : [0,1]^d \to \mathbb{T}^d, \quad (x_1, \ldots, x_d) \mapsto [(x_1, \ldots, x_d)]$$

is surjective and allows us to define a Lebesgue probability measure μ on the d-torus, through the following formula:

$$\mu(B) = m\big(\phi^{-1}(B)\big) \quad \text{for every } B \subset \mathbb{T}^d \text{ such that } \phi^{-1}(B) \text{ is measurable.}$$

This measure μ is invariant under R_θ for every θ.

We say that a vector $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$ is *rationally independent* if, for any integer numbers $n_0, n_1, \ldots, n_d$, we have that

$$n_0 + n_1\theta_1 + \cdots + n_d\theta_d = 0 \quad \Rightarrow \quad n_0 = n_1 = \cdots = n_d = 0.$$

Otherwise, we say that θ is rationally dependent. One can show that θ is rationally independent if and only if the rotation R_θ is *minimal*, meaning that the orbit $\mathcal{O}([x]) = \{R_\theta^n([x]) : n \in \mathbb{N}\}$ of every $[x] \in \mathbb{T}^d$ is a dense subset of $\mathbb{T}^d$. In this regard, see Exercises 1.3.9–1.3.10 and also Corollary 4.2.3.

1.3.5 Conservative maps

Let M be an open subset of the Euclidian space $\mathbb{R}^d$ and $f : M \to M$ be a C^1 diffeomorphism. This means that f is a bijection, both f and its inverse f^{-1} are differentiable and the two derivatives are continuous. Denote by vol the restriction to M of the Lebesgue measure (volume measure) on $\mathbb{R}^d$. The formula of change of variables asserts that, for any measurable set $B \subset M$,

$$\operatorname{vol}(f(B)) = \int_B |\det Df|\, dx. \tag{1.3.12}$$

One can easily deduce the following consequence:

Lemma 1.3.5. *A C^1 diffeomorphism $f : M \to M$ preserves the volume measure* vol *if and only if the absolute value $|\det Df|$ of its Jacobian is equal to* 1 *at every point.*

Proof. Suppose that the absolute value $|\det Df|$ of its Jacobian is equal to 1 at every point. Let E be any measurable set E and $B = f^{-1}(E)$. The formula (1.3.12) yields

$$\operatorname{vol}(E) = \int_B 1\,dx = \operatorname{vol}(B) = \operatorname{vol}(f^{-1}(E)).$$

This means that f preserves the measure vol and so we proved the "if" part of the statement.

To prove the "only if," suppose that $|\det Df(x)| > 1$ for some point $x \in M$. Then, since the Jacobian is continuous, there exists a neighborhood U of x and some number $\sigma > 1$ such that

$$|\det Df(y)| \geq \sigma \quad \text{for all } y \in U.$$

Then, applying (1.3.12) to $B = U$, we get that

$$\operatorname{vol}(f(U)) \geq \int_U \sigma\,dx \geq \sigma \operatorname{vol}(U).$$

Denote $E = f(U)$. Since $\operatorname{vol}(U) > 0$, the previous inequality implies that $\operatorname{vol}(E) > \operatorname{vol}(f^{-1}(E))$. Hence, f does not leave vol invariant. In precisely the same way, one shows that if $|\det Df(x)| < 1$ for some point $x \in M$ then f does not leave the measure vol invariant. □

1.3.6 Conservative flows

Now we discuss the invariance of the volume measure in the setting of flows $f^t : M \to M$, $t \in \mathbb{R}$. As before, take M to be an open subset of the Euclidean space $\mathbb{R}^d$. Let us suppose that the flow is C^1, in the sense that the map $(t,x) \mapsto f^t(x)$ is differentiable and all the derivatives are continuous. Then, in particular, every flow transformation $f^t : M \to M$ is a C^1 diffeomorphism: the inverse is f^{-t}. Since f^0 is the identity map and the Jacobian varies continuously, we have that $\det Df^t(x) > 0$ at every point.

Applying Lemma 1.3.5 in this context, we find that the flow preserves the volume measure if and only if

$$\det Df^t(x) = 1 \quad \text{for every } x \in U \text{ and every } t \in \mathbb{R}. \tag{1.3.13}$$

However, this is not very useful in practice because most of the time we do not have an explicit expression for f^t and, hence, it is not clear how to check the condition (1.3.13). Fortunately, there is a reasonably explicit expression for the Jacobian of the flow that can be used in some interesting situations. Let us explain this.

Let us suppose that the flow $f^t : M \to M$ corresponds to the trajectories of a C^1 vector field $F : M \to \mathbb{R}^d$. In other words, each $t \mapsto f^t(x)$ is the solution of the differential equation

$$\frac{dy}{dt} = F(y) \tag{1.3.14}$$

that has x as the initial condition (when dealing with differential equations we always assume that their solutions are defined for all times).

The *Liouville formula* relates the Jacobian of f^t to the *divergence* $\operatorname{div} F$ of the vector field:

$$\det Df^t(x) = \exp\big(\int_0^t \operatorname{div} F(f^s(x))\, ds\big) \quad \text{for every } x \text{ and every } t.$$

Recall that the divergence of a vector field F is the trace of its Jacobian matrix, that is

$$\operatorname{div} F = \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_d}{\partial x_d}. \tag{1.3.15}$$

Combining the Liouville formula with (1.3.13), we obtain:

Lemma 1.3.6 (Liouville)**.** *The flow $(f^t)_t$ associated with a C^1 vector field F preserves the volume measure if and only if the divergence of F is identically zero.*

We can extend this discussion to the case when M is any Riemannian manifold of dimension $d \geq 2$. The reader who is unfamiliar with this notion may wish to check Appendix A.4.5 before proceeding.

For simplicity, let us suppose that the manifold is orientable. Then the volume measure on M is given by a differentiable d-form ω, called the *volume form* (this remains true in the non-orientable case, except that the form ω is defined up to sign only). What this means is that the volume of any measurable set B contained in the domain of local coordinates $(x_1, \ldots, x_d)$ is given by

$$\operatorname{vol}(B) = \int_B \rho(x_1, \ldots, x_d)\, dx_1 \cdots dx_d,$$

where $\omega = \rho dx_1 \cdots dx_d$ is the expression of the volume form in those local coordinates. Let F be a C^1 vector field on M. Writing

$$F(x_1, \ldots, x_d) = (F_1(x_1, \ldots, x_d), \ldots, F_d(x_1, \ldots, x_d)),$$

we may express the divergence as

$$\operatorname{div} F = \frac{\partial(\rho F)}{\partial x_1} + \cdots + \frac{\partial(\rho F)}{\partial x_d}$$

(it can be shown that the right-hand side does not depend on the choice of the local coordinates). A proof of the following generalization of Lemma 1.3.6 can be found in Sternberg [Ste58]:

Theorem 1.3.7 (Liouville)**.** *The flow* $(f^t)_t$ *associated with a* C^1 *vector field* F *on a Riemannian manifold preserves the volume measure on the manifold if and only if* $\operatorname{div} F = 0$ *at every point.*

Then, it follows from the recurrence theorem for flows that, assuming that the manifold has finite volume (for example, if M is compact) and $\operatorname{div} F = 0$, then almost every point is recurrent for the flow of F.

1.3.7 Exercises

1.3.1. Use Lemma 1.3.3 to give another proof of the fact that the decimal expansion transformation $f(x) = 10x - [10x]$ preserves the Lebesgue measure on the interval.

1.3.2. Prove that, for any number $x \in [0,1]$ whose decimal expansion contains the block 617 (for instance, $x = 0.3375\mathbf{617}264\cdots$), that block occurs infinitely many times in the decimal expansion of x. Even more, the block 617 occurs infinitely many times in the decimal expansion of almost every $x \in [0,1]$.

1.3.3. Prove that the number 617 appears infinitely many times in the continued fraction expression of almost every number $x_0 \in (1/618, 1/617)$, that is, one has $a_n = 617$ for infinitely many values of $n \in \mathbb{N}$.

1.3.4. Let G be the Gauss map. Show that a number $x \in (0,1)$ is rational if and only if there exists $n \geq 1$ such that $G^n(x) = 0$.

1.3.5. Consider the sequence $1, 2, 4, 8, \ldots, a_n = 2^n, \ldots$ of all the powers of 2. Prove that, given any digit $i \in \{1, \ldots, 9\}$, there exist infinitely many values of n for which a_n starts with that digit.

1.3.6. Prove the following extension of Lemma 1.3.3. Let $f : M \to M$ be a C^1 local diffeomorphism on a compact Riemannian manifold M. Let vol be the volume measure on M and $\rho : M \to [0, \infty)$ be a continuous function. Then f preserves the measure $\mu = \rho \operatorname{vol}$ if and only if

$$\sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|} = \rho(y) \quad \text{for every } y \in M.$$

When f is invertible this means that f preserves the measure μ if and only if $\rho(x) = \rho(f(x))|\det Df(x)|$ for every $x \in M$.

1.3.7. Check that if A is a $d \times d$ matrix with integer coefficients and determinant different from zero then the transformation $f_A : \mathbb{T}^d \to \mathbb{T}^d$ defined on the torus by $f_A([x]) = [A(x)]$ preserves the Lebesgue measure on $\mathbb{T}^d$.

1.3.8. Show that the Lebesgue measure on S^1 is the only probability measure invariant under all the rotations of S^1, even if we restrict to *rational* rotations. [Note: We will see in Chapter 6 that, for any *irrational* θ, the Lebesgue measure is the unique probability measure invariant under R_θ.]

1.3.9. Suppose that $\theta = (\theta_1, \dots, \theta_d)$ is rationally dependent. Show that there exists a continuous non-constant function $\varphi : \mathbb{T}^d \to \mathbb{C}$ such that $\varphi \circ R_\theta = \varphi$. Conclude that there exist non-empty open subsets U and V of $\mathbb{T}^d$ that are disjoint and invariant under R_θ, in the sense that $R_\theta(U) = U$ and $R_\theta(V) = V$. Deduce that no orbit $\mathcal{O}([x])$ of the rotation R_θ is dense in $\mathbb{T}^d$.

1.3.10. Suppose that $\theta = (\theta_1, \dots, \theta_d)$ is rationally independent. Prove that if V is a non-empty open subset of $\mathbb{T}^d$ invariant under R_θ, then V is dense in $\mathbb{T}^d$. Conclude that $\cup_{n\in\mathbb{Z}} R_\theta^n(U)$ is dense in the torus, for every non-empty open subset U. Deduce that there exists $[x]$ whose orbit $\mathcal{O}([x])$ under the rotation R_θ is dense in $\mathbb{T}^d$. Conclude that $\mathcal{O}([y])$ is dense in $\mathbb{T}^d$ for every $[y]$.

1.3.11. Let U be an open subset of $\mathbb{R}^{2d}$ and $H : U \to \mathbb{R}$ be a C^2 function. Denote by $(p_1, \dots, p_d, q_1, \dots, q_d)$ the coordinate variables in $\mathbb{R}^{2d}$. The *Hamiltonian vector field* associated with H is defined by

$$F(p_1, \dots, p_d, q_1, \dots, q_d) = \left(\frac{\partial H}{\partial q_1}, \dots, \frac{\partial H}{\partial q_d}, -\frac{\partial H}{\partial p_1}, \dots, -\frac{\partial H}{\partial p_d}\right).$$

Check that the flow defined by F preserves the volume measure.

1.3.12. Let $f : U \to U$ be a C^1 diffeomorphism preserving the volume measure on an open subset U of $\mathbb{R}^d$. Let $H : U \to \mathbb{R}$ be a first integral of f, that is, a C^1 function such that $H \circ f = H$. Let c be a regular value of H and ds be the volume measure defined on the hypersurface $H_c = H^{-1}(c)$ by the restriction of the Riemannian metric of $\mathbb{R}^d$. Prove that the restriction of f to the hypersurface H_c preserves the measure $ds/\|\operatorname{grad} H\|$.

1.4 Induction

In this section we describe a general method, based on the Poincaré recurrence theorem, to construct from a given system (f, μ) other systems, that we refer to as *systems induced* by (f, μ). The reason this is interesting is the following. On the one hand, it is often the case that an induced system is easier to analyze, because it has better global properties than the original one. On the other hand, interesting conclusions about the original system can often be obtained from analyzing the induced one. Examples will appear in a while.

1.4.1 First return map

Let $f : M \to M$ be a measurable transformation and μ be an invariant probability measure. Let $E \subset M$ be a measurable set with $\mu(E) > 0$ and $\rho(x) = \rho_E(x)$ be the first-return time of x to E, as given by (1.2.1). The *first-return map* to the domain E is the map g given by

$$g(x) = f^{\rho(x)}(x)$$

whenever $\rho(x)$ is finite. The Poincaré recurrence theorem ensures that this is the case for μ-almost every $x \in E$ and so g is defined on a full measure subset of E. We also denote by μ_E the restriction of μ to the measurable subsets E.

Proposition 1.4.1. *The measure μ_E is invariant under the map $g : E \to E$.*

Proof. For every $k \geq 1$, denote by E_k the subset of points $x \in E$ such that $\rho(x) = k$. By definition, $g(x) = f^k(x)$ for every $x \in E_k$. Let B be any measurable subset of E. Then

$$\mu(g^{-1}(B)) = \sum_{k=1}^{\infty} \mu(f^{-k}(B) \cap E_k). \tag{1.4.1}$$

On the other hand, since μ is f-invariant,

$$\mu\big(B\big) = \mu\big(f^{-1}(B)\big) = \mu\big(f^{-1}(B) \cap E_1\big) + \mu\big(f^{-1}(B) \setminus E\big). \tag{1.4.2}$$

Analogously,

$$\begin{aligned}\mu\big(f^{-1}(B) \setminus E\big) &= \mu\big(f^{-2}(B) \setminus f^{-1}(E)\big)\\ &= \mu\big(f^{-2}(B) \cap E_2\big) + \mu\big(f^{-2}(B) \setminus (E \cup f^{-1}(E))\big).\end{aligned}$$

Replacing this expression in (1.4.2), we find that

$$\mu\big(B\big) = \sum_{k=1}^{2} \mu\big(f^{-k}(B) \cap E_k\big) + \mu\Big(f^{-2}(B) \setminus \bigcup_{k=0}^{1} f^{-k}(E)\Big).$$

Repeating this argument successively, we obtain

$$\mu\big(B\big) = \sum_{k=1}^{n} \mu\big(f^{-k}(B) \cap E_k\big) + \mu\Big(f^{-n}(B) \setminus \bigcup_{k=0}^{n-1} f^{-k}(E)\Big). \tag{1.4.3}$$

Now let us go to the limit when $n \to \infty$. It is clear that the last term is bounded above by $\mu\big(f^{-n}(E) \setminus \bigcup_{k=0}^{n-1} f^{-k}(E)\big)$. So, using Remark 1.2.3, that term converges to zero when $n \to \infty$. In this way we conclude that

$$\mu\big(B\big) = \sum_{k=1}^{\infty} \mu\big(f^{-k}(B) \cap E_k\big).$$

Together with (1.4.1), this shows that $\mu(g^{-1}(B)) = \mu(B)$ for every measurable subset B of E. That is to say, the measure μ_E is invariant under g. □

Example 1.4.2. Consider the transformation $f : [0, \infty) \to [0, \infty)$ defined by

$$f(0) = 0 \quad \text{and} \quad f(x) = 1/x \text{ if } x \in (0, 1) \quad \text{and} \quad f(x) = x - 1 \text{ if } x \geq 1.$$

Let $E = [0, 1]$. The time ρ of first return to E is given by

$$\rho(0) = 1 \quad \text{and} \quad \rho(x) = k + 1 \text{ if } x \in (1/(k + 1), 1/k] \text{ with } k \geq 1.$$

So, the first-return map to E is given by

$$g(0) = 0 \quad \text{and} \quad g(x) = 1/x - k \text{ if } x \in (1/(k+1), 1/k] \text{ with } k \geq 1.$$

In other words, g is the Gauss map. We saw in Section 1.3.2 that the Gauss map admits an invariant probability measure equivalent to the Lebesgue measure on $[0,1)$. From this, one can draw some interesting conclusions about the original map f. For instance, using the ideas in the next section one finds that f admits an (infinite) invariant measure equivalent to the Lebesgue measure on $[0,\infty)$.

1.4.2 Induced transformations

In an opposite direction, given any measure ν invariant under $g : E \to E$, we may construct a certain related measure ν_ρ that is invariant under $f : M \to M$. For this, g does not even have to be a first-return map: the construction that we present below is valid for any map *induced* from f, that is, any map of the form

$$g : E \to E, \quad g(x) = f^{\rho(x)}(x), \tag{1.4.4}$$

where $\rho : E \to \mathbb{N}$ is a measurable function (it suffices that ρ is defined on some full measure subset of E). As before, we denote by E_k the subset of points $x \in E$ such that $\rho(x) = k$. Then we define

$$\nu_\rho(B) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(f^{-n}(B) \cap E_k), \tag{1.4.5}$$

for every measurable set $B \subset M$.

Proposition 1.4.3. *The measure ν_ρ defined in* (1.4.5) *is invariant under f and satisfies $\nu_\rho(M) = \int_E \rho \, d\nu$. In particular, ν_ρ is finite if and only if the function ρ is integrable with respect to ν.*

Proof. First, let us prove that ν_ρ is invariant. By the definition (1.4.5),

$$\nu_\rho\big(f^{-1}(B)\big) = \sum_{n=0}^{\infty} \sum_{k>n} \nu\big(f^{-(n+1)}(B) \cap E_k\big) = \sum_{n=1}^{\infty} \sum_{k\geq n} \nu\big(f^{-n}(B) \cap E_k\big).$$

We may rewrite this expression as follows:

$$\nu_\rho\big(f^{-1}(B)\big) = \sum_{n=1}^{\infty} \sum_{k>n} \nu\big(f^{-n}(B) \cap E_k\big) + \sum_{k=1}^{\infty} \nu\big(f^{-k}(B) \cap E_k\big). \tag{1.4.6}$$

Concerning the last term, observe that

$$\sum_{k=1}^{\infty} \nu\big(f^{-k}(B) \cap E_k\big) = \nu\big(g^{-1}(B)\big) = \nu\big(B\big) = \sum_{k=1}^{\infty} \nu\big(B \cap E_k\big),$$

since ν is invariant under g. Replacing this in (1.4.6), we see that

$$\nu_\rho\big(f^{-1}(B)\big) = \sum_{n=1}^{\infty}\sum_{k>n}\nu\big(f^{-n}(B)\cap E_k\big) + \sum_{k=1}^{\infty}\nu\big(B\cap E_k\big) = \nu_\rho\big(B\big)$$

for every measurable set $B \subset E$. The second claim is a direct consequence of the definitions:

$$\nu_\rho(M) = \sum_{n=0}^{\infty}\sum_{k>n}\nu(f^{-n}(M)\cap E_k) = \sum_{n=0}^{\infty}\sum_{k>n}\nu(E_k) = \sum_{k=1}^{\infty}k\nu(E_k) = \int_E \rho\,d\nu.$$

This completes the proof. □

It is interesting to analyze how this construction relates to the one in the previous section when g is a first-return map of f and the measure ν is the restriction $\mu \mid E$ of some invariant measure μ of f:

Corollary 1.4.4. *If g is the first-return map of f to a measurable subset E and $\nu = \mu \mid E$, then*

1. *$\nu_\rho(B) = \nu(B) = \mu(B)$ for every measurable set $B \subset E$.*

2. *$\nu_\rho(B) \le \mu(B)$ for every measurable set $B \subset M$.*

Proof. By definition, $f^{-n}(E)\cap E_k = \emptyset$ for every $0 < n < k$. This implies that, given any measurable set $B \subset E$, all the terms with $n > 0$ in the definition (1.4.5) are zero. Hence, $\nu_\rho(B) = \sum_{k>0}\nu(B\cap E_k) = \nu(B)$ as claimed in the first part of the statement.

Consider any measurable set $B \subset M$. Then,

$$\begin{aligned}\mu\big(B\big) &= \mu\big(B\cap E\big) + \mu\big(B\cap E^c\big) = \nu(B\cap E) + \mu\big(B\cap E^c\big)\\ &= \sum_{k=1}^{\infty}\nu\big(B\cap E_k\big) + \mu\big(B\cap E^c\big).\end{aligned} \tag{1.4.7}$$

Since μ is invariant, $\mu(B\cap E^c) = \mu\big(f^{-1}(B)\cap f^{-1}(E^c)\big)$. Then, as in the previous equality,

$$\begin{aligned}\mu\big(B\cap E^c\big) &= \mu\big(f^{-1}(B)\cap E\cap f^{-1}(E^c)\big) + \mu\big(f^{-1}(B)\cap E^c\cap f^{-1}(E^c)\big)\\ &= \sum_{k=2}^{\infty}\nu\big(f^{-1}(B)\cap E_k\big) + \mu\big(f^{-1}(B)\cap E^c\cap f^{-1}(E^c)\big).\end{aligned}$$

Replacing this in (1.4.7), we find that

$$\mu\big(B\big) = \sum_{n=0}^{1}\sum_{k>n}\nu\big(f^{-n}(B)\cap E_k\big) + \mu\big(f^{-1}(B)\cap\bigcap_{n=0}^{1}f^{-n}(E^c)\big).$$

Repeating this argument successively, we get that

$$\begin{aligned}\mu(B) &= \sum_{n=0}^{N}\sum_{k>n}\nu\big(f^{-n}(B)\cap E_k\big) + \mu\big(f^{-N}(B)\cap\bigcap_{k=0}^{N} f^{-n}(E^c)\big)\\ &\geq \sum_{n=0}^{N}\sum_{k>n}\nu\big(f^{-n}(B)\cap E_k\big) \quad \text{for every } N\geq 1.\end{aligned}$$

Taking the limit as $N\to\infty$, we conclude that $\mu(B)\geq\nu_\rho(B)$. □

We also have from the Kač theorem (Theorem 1.2.2) that

$$\nu_\rho(M) = \int_E \rho\,d\nu = \int_E \rho\,d\mu = \mu(M) - \mu(E_0^*).$$

So, it follows from Corollary 1.4.4 that $\nu_\rho = \mu$ if and only if $\mu(E_0^*) = 0$.

Example 1.4.5 (Manneville-Pomeau)**.** Given $d>0$, let a be the only number in $(0,1)$ such that $a(1+a^d)=1$. Then define $f:[0,1]\to[0,1]$ as follows:

$$f(x) = x(1+x^d) \quad\text{if}\quad x\in[0,a] \quad\text{and}\quad f(x) = \frac{x-a}{1-a} \quad\text{if}\quad x\in(a,1].$$

The graph of f is depicted on the left-hand side of Figure 1.3. Observe that $|f'(x)|\geq 1$ at every point, and the inequality is strict at every $x>0$. Let $(a_n)_n$ be the sequence on the interval $[0,a]$ defined by $a_1=a$ and $f(a_{n+1})=a_n$ for $n\geq 1$. We also write $a_0=1$. Some properties of this sequence are studied in Exercise 1.4.2.

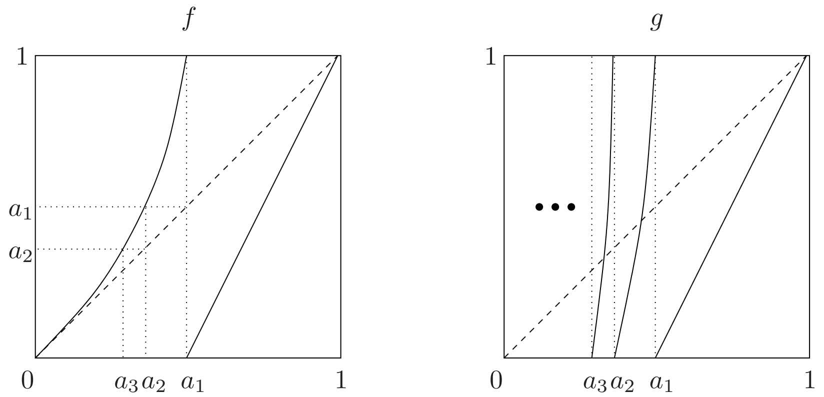

Figure 1.3: Construction of an induced transformation

Now consider the map $g(x) = f^{\rho(x)}(x)$, where

$$\rho:[0,1]\to\mathbb{N}, \quad \rho(x) = 1+\min\{n\geq 0 : f^n(x)\in(a,1]\}.$$

In other words, $\rho(x) = k$ and so $g(x) = f^k(x)$ for every $x \in (a_k, a_{k-1}]$. The graph of g is represented on the right-hand side of Figure 1.3. Note that the restriction to each interval $(a_k, a_{k-1}]$ is a bijection onto $(0, 1]$. A key point is that the induced map g is *expanding*:

$$|g'(x)| \geq \frac{1}{1-a} > 1 \quad \text{for every } x \in [0,1].$$

Using the ideas that will be developed in Chapter 11, one can show that g admits a unique invariant probability measure ν equivalent to the Lebesgue measure on $(0, 1]$. In fact, the density (Radon-Nikodym derivative) of ν with respect to the Lebesgue measure is bounded from zero and infinity. Then, the f-invariant measure ν_ρ in (1.4.5) is equivalent to Lebesgue measure. It follows (see Exercise 1.4.2) that this measure is finite if and only if $d \in (0, 1)$.

1.4.3 Kakutani-Rokhlin towers

It is possible, and useful, to generalize the previous constructions even further, by omitting the initial transformation $f : M \to M$ altogether. More precisely, given a transformation $g : E \to E$, a measure ν on E invariant under g and a measurable function $\rho : E \to \mathbb{N}$, we are going to *construct* a transformation $f : M \to M$ and a measure ν_ρ invariant under f such that E can be identified with a subset of M, g is the first-return map of f to E, with first-return time given by ρ, and the restriction of ν_ρ to E coincides with ν.

This transformation f is called the *Kakutani-Rokhlin tower* of g with time ρ. The measure ν_ρ is finite if and only if ρ is integrable with respect to ν. They are constructed as follows. Begin by defining

$$\begin{aligned} M &= \{(x, n) : x \in E \text{ and } 0 \leq n < \rho(x)\} \\ &= \bigcup_{k=1}^{\infty} \bigcup_{n=0}^{k-1} E_k \times \{n\}. \end{aligned}$$

In other words, M consists of k copies of each set $E_k = \{x \in E : \rho(x) = k\}$, "piled up" on top of each other. We call each $\cup_{k>n} E_k \times \{n\}$ the *n-th floor* of M. See Figure 1.4.

Next, define $f : M \to M$ as follows:

$$f(x, n) = \begin{cases} (x, n+1) & \text{if } n < \rho(x) - 1 \\ (g(x), 0) & \text{if } n = \rho(x) - 1 \end{cases}.$$

In other words, the dynamics "lifts" each point (x, n) one floor at a time, until reaching the floor $\rho(x) - 1$; at that stage, the point "falls" directly to $(g(x), 0)$ on the ground (zero-th) floor. The ground floor $E \times \{0\}$ is naturally identified with the set E. Besides, the first-return map to $E \times \{0\}$ corresponds precisely to $g : E \to E$.

Finally, the measure ν_ρ is defined by

$$\nu_\rho \mid (E_k \times \{n\}) = \nu \mid E_k$$

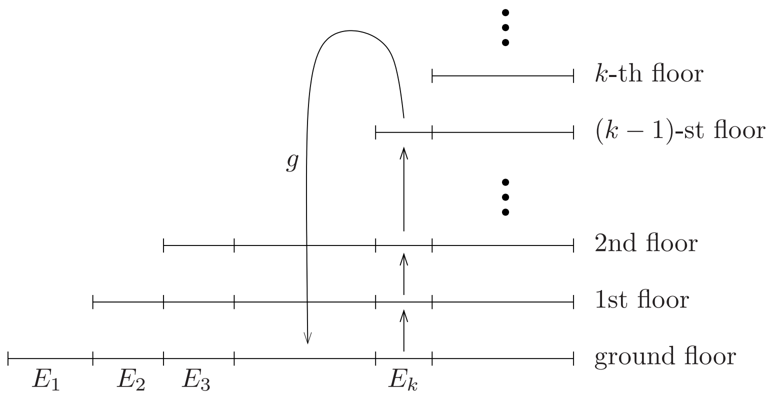

Figure 1.4: Kakutani-Rokhlin tower of g with time ρ

for every $0 \le n < k$. It is clear that the restriction of ν_ρ to the ground floor coincides with ν. Moreover, ν_ρ is invariant under f and

$$\nu_\rho(M) = \sum_{k=1}^{\infty} k\nu(E_k) = \int_E \rho \, d\nu.$$

This completes the construction of the Kakutani-Rokhlin tower.

1.4.4 Exercises

1.4.1. Let $f : S^1 \to S^1$ be the transformation $f(x) = 2x \mod \mathbb{Z}$. Show that the function $\tau(x) = \min\{k \ge 0 : f^k(x) \in (1/2, 1)\}$ is integrable with respect to the Lebesgue measure. State and prove a corresponding result for any C^1 transformation $g : S^1 \to S^1$ that is close to f, in the sense that $\sup_x\{\|g(x) - f(x)\|, \|g'(x) - f'(x)\|\}$ is sufficiently small.

1.4.2. Consider the measure ν_ρ and the sequence $(a_n)_n$ defined in Example 1.4.5. Check that ν_ρ is always σ-finite. Show that $(a_n)_n$ is decreasing and converges to zero. Moreover, there exist $c_1, c_2, c_3, c_4 > 0$ such that

$$c_1 \le a_j j^{1/d} \le c_2 \quad \text{and} \quad c_3 \le \big(a_j - a_{j+1}\big) j^{1+1/d} \le c_4 \quad \text{for every } j. \qquad (1.4.8)$$

Deduce that the g-invariant measure ν_ρ is finite if and only if $d \in (0, 1)$.

1.4.3. Let $\sigma : \Sigma \to \Sigma$ be the map defined on the space $\Sigma = \{1, \dots, d\}^{\mathbb{Z}}$ by $\sigma((x_n)_n) = (x_{n+1})_n$. Describe the first-return map g to the subset $\{(x_n)_n \in \Sigma : x_0 = 1\}$.

1.4.4. [Kakutani-Rokhlin lemma] Let $f : M \to M$ be an invertible transformation and μ be an invariant probability measure without atoms and such that $\mu(\cup_{n\in\mathbb{N}} f^n(E)) = 1$ for every $E \subset M$ with $\mu(E) > 0$. Show that for every

$n \geq 1$ and $\varepsilon > 0$ there exists a measurable set $B \subset M$ such that the iterates $B, f(B), \ldots, f^{n-1}(B)$ are pairwise disjoint and the complement of their union has measure less than ε. In particular, this holds for every invertible system that is *aperiodic*, that is, whose periodic points form a zero measure set.

1.4.5. Let $f : M \to M$ be a transformation and $(H_j)_{j\geq 1}$ be a collection of subsets of M such that if $x \in H_n$ then $f^j(x) \in H_{n-j}$ for every $0 \leq j < n$. Let H be the set of points that belong to H_j for infinitely many values of j, that is, $H = \cap_{k=1}^{\infty} \cup_{j=k}^{\infty} H_j$. For $y \in H$, define $\tau(y) = \min\{j \geq 1 : y \in H_j\}$ and $T(y) = f^{\tau(y)}(y)$. Observe that T maps H inside H. Moreover, show that

$$\limsup_n \frac{1}{n}\#\{1 \leq j \leq n : x \in H_j\} \geq \theta > 0 \quad \Rightarrow \quad \liminf_k \frac{1}{k}\sum_{i=0}^{k-1} \tau(T^i(x)) \leq \frac{1}{\theta}.$$

1.4.6. Let $f : M \to M$ be a transformation preserving a measure μ. Let $(H_j)_{j\geq 1}$ and $\tau : M \to \mathbb{N}$ be as in Exercise 1.4.5. Consider the sequence of functions $(\tau_n)_n$ defined by $\tau_1(x) = \tau(x)$ and $\tau_n(x) = \tau(f^{\tau_{n-1}(x)}(x)) + \tau_{n-1}(x)$ for $n > 1$. Suppose that

$$\limsup_n \frac{1}{n}\#\{1 \leq j \leq n : x \in H_j\} \geq \theta > 0 \quad \text{for } \mu\text{-almost every } x \in M.$$

Show that $\tau_{n+1}(x)/\tau_n(x) \to 1$ for μ-almost every $x \in M$. [Note: Sequences with this property are called *non-lacunary*.]

1.5 Multiple recurrence theorems

Now we consider finite families of *commuting maps* $f_i : M \to M$, $i = 1, \ldots, q$, that is, such that

$$f_i \circ f_j = f_j \circ f_i \quad \text{for every } i, j \in \{1, \ldots, q\}.$$

Our goal is to explain that the results in Section 1.2 extend to this setting: we find points that are *simultaneously recurrent* for these transformations.

The first result in this direction generalizes the Birkhoff recurrence theorem (Theorem 1.2.6):

Theorem 1.5.1 (Birkhoff multiple recurrence)**.** *Let M be a compact metric space and $f_1, \ldots, f_q : M \to M$ be continuous commuting maps. Then there exists $a \in M$ and a sequence $(n_k)_k \to \infty$ such that*

$$\lim_k f_i^{n_k}(a) = a \quad \text{for every } i = 1, \ldots, q. \tag{1.5.1}$$

The key point here is that the sequence $(n_k)_k$ does not depend on i: we say that the point a is *simultaneously recurrent* for all the maps f_i, $i = 1, \ldots, q$. A proof of Theorem 1.5.1 is given in Section 1.5.1. Next, we discuss the following generalization of the Poincaré recurrence theorem (Theorem 1.2.1):

Theorem 1.5.2 (Poincaré multiple recurrence). *Let $(M, \mathcal{B}, \mu)$ be a probability space and $f_i : M \to M$, $i = 1, \ldots, q$ be measurable commuting maps that preserve the measure μ. Then, given any set $E \subset M$ with positive measure, there exists $n \geq 1$ such that*

$$\mu\big(E \cap f_1^{-n}(E) \cap \cdots \cap f_q^{-n}(E)\big) > 0.$$

In other words, for a positive measure subset of points $x \in E$, their orbits under all the maps f_i, $i = 1, \ldots, q$ return to E *simultaneously* at time n (we say that n is a *simultaneous return* of x to E): once more, the crucial point with the statement is that n does not depend on i.

The proof of Theorem 1.5.2 will not be presented here; we refer the interested reader to the book of Furstenberg [Fur77]. We are just going to mention some direct consequences and, in Chapter 2, we will use this theorem to prove the Szemerédi theorem on the existence of arithmetic progressions inside "dense" subsets of integer numbers.

To begin with, observe that the set of simultaneous returns is always infinite. Indeed, let n be as in the statement of Theorem 1.5.2. Applying the theorem to the set $F = E \cap f_1^{-n}(E) \cap \cdots \cap f_q^{-n}(E)$, we find $m \geq 1$ such that

$$\mu\big(E \cap f_1^{-(m+n)}(E) \cap \cdots \cap f_q^{-(m+n)}(E)\big) \geq \mu\big(F \cap f_1^{-m}(F) \cap \cdots \cap f_q^{-m}(F)\big) > 0.$$

Thus, $m+n$ is also a simultaneous return to E, for all the points in some subset of E with positive measure.

It follows that, for any set $E \subset M$ with $\mu(E) > 0$ and for μ-almost every point $x \in E$, there exist infinitely many *simultaneous returns* of x to E. Indeed, suppose there is a positive measure set $F \subset E$ such that every point of F has a finite number of simultaneous returns to E. On the one hand, up to replacing F by a suitable subset, we may suppose that the simultaneous returns to E of all the points of F are bounded by some $k \geq 1$. On the other hand, using the previous paragraph, there exists $n > k$ such that $G = F \cap f_1^{-n}(F) \cap \cdots \cap f_q^{-n}(F)$ has positive measure. Now, it is clear from the definition that n is a simultaneous return to E of every $x \in G$. This contradicts the choice of F, thus proving our claim.

Another direct corollary is the Birkhoff multiple recurrence theorem (Theorem 1.5.1). Indeed, if $f_i : M \to M$, $i = 1, \ldots, q$ are continuous commuting transformations on a compact metric space then there exists some probability measure μ that is invariant under all these transformations (this fact will be checked in the next chapter, see Exercise 2.2.2). From this point on, we may argue exactly as in the proof of Theorem 1.2.4. More precisely, consider any countable basis $\{U_k\}$ for the topology of M. According to the previous paragraph, for every k there exists a set $\tilde{U}_k \subset U_k$ with zero measure such that every point in $U_k \setminus \tilde{U}_k$ has infinitely many simultaneous returns to U_k. Then $\tilde{U} = \cup_k \tilde{U}_k$ has measure zero and every point in its complement is simultaneously recurrent, in the sense of Theorem 1.5.1.

1.5.1 Birkhoff multiple recurrence theorem

In this section we prove Theorem 1.5.1 in the case when the transformations $f_1, \dots, f_q$ are homeomorphisms of M, which suffices for all our purposes in the present chapter. The general case may be deduced easily (see Exercise 2.4.7) using the concept of natural extension, which we will present in the next chapter.

The theorem may be reformulated in the following useful way. Consider the transformation $F : M^q \to M^q$ defined on the product space $M^q = M \times \cdots \times M$ by $F(x_1, \dots, x_q) = (f_1(x_1), \dots, f_q(x_q))$. Denote by Δ_q the *diagonal* of M^q, that is, the subset of points of the form $\tilde{x} = (x, \dots, x)$. Theorem 1.5.1 claims, precisely, that there exist $\tilde{a} \in \Delta_q$ and $(n_k)_k \to \infty$ such that

$$\lim_k F^{n_k}(\tilde{a}) = \tilde{a}. \tag{1.5.2}$$

The proof of Theorem 1.5.1 is by induction on the number q of transformations. The case $q = 1$ is contained in Theorem 1.2.6. Consider any $q \geq 2$ and suppose that the statement is true for every family of $q - 1$ commuting homeomorphisms. We are going to prove that it is true for the family $f_1, \dots, f_q$.

Let $\mathcal{G}$ be the (abelian) group generated by the homeomorphisms $f_1, \dots, f_q$. We say that a set $X \subset M$ is $\mathcal{G}$*-invariant* if $g(X) \subset X$ for every $g \in \mathcal{G}$. Observing that the inverse g^{-1} is also in $\mathcal{G}$, we see that this implies $g(X) = X$ for every $g \in \mathcal{G}$. Just as we did in Theorem 1.2.6, we may use Zorn's lemma to conclude that there exists some minimal, non-empty, closed, $\mathcal{G}$-invariant set $X \subset M$ (this is Exercise 1.5.2). The statement of the theorem is not affected if we replace M by X. Thus, it is no restriction to assume that the ambient space M is minimal. This assumption is used as follows:

Lemma 1.5.3. *If M is minimal then for every non-empty open set $U \subset M$ there exists a finite subset $\mathcal{H} \subset \mathcal{G}$ such that*

$$\bigcup_{h \in \mathcal{H}} h^{-1}(U) = M.$$

Proof. For any $x \in M$, the closure of the orbit $\mathcal{G}(x) = \{g(x) : g \in \mathcal{G}\}$ is a non-empty, closed, $\mathcal{G}$-invariant subset of M. So, the hypothesis that M is minimal implies that every orbit $\mathcal{G}(x)$ is dense in M. In particular, there is $g \in \mathcal{G}$ such that $g(x) \in U$. This proves that $\{g^{-1}(U) : g \in \mathcal{G}\}$ is an open cover of M. By compactness, it follows that there exists a finite subcover, as claimed. □

Consider the product M^q endowed with the distance function

$$d\big((x_1, \dots, x_q), (y_1, \dots, y_q)\big) = \max\{d(x_i, y_i) : 1 \leq i \leq q\}.$$

Note that the map $M \to \Delta_q$, $x \mapsto \tilde{x} = (x, \dots, x)$ is a homeomorphism, and even an isometry for this choice of a distance. Every open set $U \subset M$ corresponds to an open set $\widetilde{U} \subset \Delta_q$ through this homeomorphism. Given any $g \in \mathcal{G}$, we denote by $\tilde{g} : M^q \to M^q$ the homeomorphism defined by $\tilde{g}(x_1, \dots, x_q) = (g(x_1), \dots, g(x_q))$. The fact that the group $\mathcal{G}$ is abelian implies that $\tilde{g}$ commutes with F; note also that every $\tilde{g}$ preserves the diagonal Δ_q. Then the conclusion of Lemma 1.5.3 may be rewritten in the following form:

$$\bigcup_{h\in\mathcal{H}} \tilde{h}^{-1}(\widetilde{U}) = \Delta_q. \tag{1.5.3}$$

Lemma 1.5.4. *Given $\varepsilon > 0$ there exist $\tilde{x} \in \Delta_q$, $\tilde{y} \in \Delta_q$ and $n \geq 1$ such that $d(F^n(\tilde{x}), \tilde{y}) < \varepsilon$.*

Proof. Define $g_i = f_i \circ f_q^{-1}$ for each $i = 1, \ldots, q-1$. Since the maps f_i commute with each other, so do the maps g_i. Then, by induction, there exist $y \in M$ and $(n_k)_k \to \infty$ such that

$$\lim_k g_i^{n_k}(y) = y \quad \text{for every } i = 1, \ldots, q-1.$$

Denote $x_k = f_q^{-n_k}(y)$ and consider $\tilde{x}_k = (x_k, \ldots, x_k) \in \Delta_q$. Then,

$$\begin{aligned} F^{n_k}(\tilde{x}_k) &= (f_1^{n_k} f_q^{-n_k}(y), \ldots, f_{q-1}^{n_k} f_q^{-n_k}(y), f_q^{n_k} f_q^{-n_k}(y)) \\ &= (g_1^{n_k}(y), \ldots, g_{q-1}^{n_k}(y), y) \end{aligned}$$

converges to $(y, \ldots, y, y)$ when $k \to \infty$. This proves the lemma, with $\tilde{x} = \tilde{x}_k$, $\tilde{y} = (y, \ldots, y, y)$ and $n = n_k$ for every k sufficiently large. ☐

The next step is to show that the point $\tilde{y}$ in Lemma 1.5.4 is arbitrary:

Lemma 1.5.5. *Given $\varepsilon > 0$ and $\tilde{z} \in \Delta_q$ there exist $\tilde{w} \in \Delta_q$ and $m \geq 1$ such that $d(F^m(\tilde{w}), \tilde{z}) < \varepsilon$.*

Proof. Given $\varepsilon > 0$ and $\tilde{z} \in \Delta_q$, consider $\widetilde{U}$ = open ball of center $\tilde{z}$ and radius $\varepsilon/2$. By Lemma 1.5.3 and the observation (1.5.3), we may find a finite set $\mathcal{H} \subset \mathcal{G}$ such that the sets $\tilde{h}^{-1}(\widetilde{U})$, $h \in \mathcal{H}$ cover Δ_q. Since the elements of $\mathcal{G}$ are (uniformly) continuous functions, there exists $\delta > 0$ such that

$$d(\tilde{x}_1, \tilde{x}_2) < \delta \quad \Rightarrow \quad d(\tilde{h}(\tilde{x}_1), \tilde{h}(\tilde{x}_2)) < \varepsilon/2 \quad \text{for every } h \in \mathcal{H}.$$

By Lemma 1.5.4 there exist $\tilde{x}, \tilde{y} \in \Delta_q$ and $n \geq 1$ such that $d(F^n(\tilde{x}), \tilde{y}) < \delta$. Fix $h \in \mathcal{H}$ such that $\tilde{y} \in \tilde{h}^{-1}(\widetilde{U})$. Then,

$$d\big(\tilde{h}(F^n(\tilde{x})), \tilde{z}\big) \leq d\big(\tilde{h}(F^n(\tilde{x})), \tilde{h}(\tilde{y})\big) + d\big(\tilde{h}(\tilde{y}), \tilde{z}\big) < \varepsilon/2 + \varepsilon/2.$$

Take $\tilde{w} = \tilde{h}(\tilde{x})$. Since $\tilde{h}$ commutes with F^n, the previous inequality implies that $d(F^n(\tilde{w}), \tilde{z}) < \varepsilon$, as we wanted to prove. ☐

Next, we prove that one may take $\tilde{x} = \tilde{y}$ in Lemma 1.5.4:

Lemma 1.5.6 (Bowen)**.** *Given $\varepsilon > 0$ there exist $\tilde{v} \in \Delta_q$ and $k \geq 1$ with $d(F^k(\tilde{v}), \tilde{v}) < \varepsilon$.*

Proof. Given $\varepsilon > 0$ and $\tilde{z}_0 \in \Delta_q$, consider the sequences ε_j, m_j and $\tilde{z}_j$, $j \geq 1$ defined by recurrence as follows. Initially, take $\varepsilon_1 = \varepsilon/2$.

- By Lemma 1.5.5 there are $\tilde{z}_1 \in \Delta_q$ and $m_1 \geq 1$ with $d(F^{m_1}(\tilde{z}_1), \tilde{z}_0) < \varepsilon_1$.
- By the continuity of F^{m_1}, there exists $\varepsilon_2 < \varepsilon_1$ such that $d(\tilde{z}, \tilde{z}_1) < \varepsilon_2$ implies $d(F^{m_1}(\tilde{z}), \tilde{z}_0) < \varepsilon_1$.

Next, given any $j \geq 2$:

- By Lemma 1.5.5 there are $\tilde{z}_j \in \Delta_q$ and $m_j \geq 1$ with $d(F^{m_j}(\tilde{z}_j), \tilde{z}_{j-1}) < \varepsilon_j$.
- By the continuity of F^{m_j}, there exists $\varepsilon_{j+1} < \varepsilon_j$ such that $d(\tilde{z}, \tilde{z}_j) < \varepsilon_{j+1}$ implies $d(F^{m_j}(\tilde{z}), \tilde{z}_{j-1}) < \varepsilon_j$.

In particular, for any $i < j$,

$$d(F^{m_{i+1}+\cdots+m_j}(\tilde{z}_j), \tilde{z}_i) < \varepsilon_{i+1} \leq \frac{\varepsilon}{2}.$$

Since Δ_q is compact, we can find i, j with $i < j$ such that $d(\tilde{z}_i, \tilde{z}_j) < \varepsilon/2$. Take $k = m_{i+1} + \cdots + m_j$. Then,

$$d(F^k(\tilde{z}_j), \tilde{z}_j) \leq d(F^k(\tilde{z}_j), \tilde{z}_i) + d(\tilde{z}_i, \tilde{z}_j) < \varepsilon.$$

This completes the proof of the lemma. □

Now we are ready to conclude the proof of Theorem 1.5.1. For that, let us consider the function

$$\phi : \Delta_q \to [0, \infty), \quad \phi(\tilde{x}) = \inf\{d(F^n(\tilde{x}), \tilde{x}) : n \geq 1\}.$$

Observe that ϕ is upper semi-continuous: given any $\varepsilon > 0$, every point $\tilde{x}$ admits some neighborhood V such that $\phi(\tilde{y}) < \phi(\tilde{x}) + \varepsilon$ for every $y \in V$. This is an immediate consequence of the fact that ϕ is given by the infimum of a family of continuous functions. Then (Exercise 1.5.4), ϕ admits some continuity point $\tilde{a}$. We are going to show that this point satisfies the conclusion of Theorem 1.5.1.

Let us begin by observing that $\phi(\tilde{a}) = 0$. Indeed, suppose that $\phi(\tilde{a})$ is positive. Then, by continuity, there exist $\beta > 0$ and a neighborhood V of $\tilde{a}$ such that $\phi(\tilde{y}) \geq \beta > 0$ for every $\tilde{y} \in V$. Then,

$$d(F^n(\tilde{y}), \tilde{y}) \geq \beta \quad \text{for every } y \in V \text{ and } n \geq 1. \tag{1.5.4}$$

On the other hand, according to (1.5.3), for every $\tilde{x} \in \Delta_q$ there exists $h \in \mathcal{H}$ such that $\tilde{h}(\tilde{x}) \in V$. Since the transformations h are uniformly continuous, we may fix $\alpha > 0$ such that

$$d(\tilde{z}, \tilde{w}) < \alpha \quad \Rightarrow \quad d\big(\tilde{h}(\tilde{z}), \tilde{h}(\tilde{w})\big) < \beta \quad \text{for every } h \in \mathcal{H}. \tag{1.5.5}$$

By Lemma 1.5.6, there exists $n \geq 1$ such that $d(\tilde{x}, F^n(\tilde{x})) < \alpha$. Then, using (1.5.5) and recalling that F commutes with every $\tilde{h}$,

$$d\big(\tilde{h}(\tilde{x}), F^n(\tilde{h}(\tilde{x}))\big) < \beta.$$

This contradicts (1.5.4). This contradiction proves that $\phi(\tilde{a}) = 0$, as claimed.

In other words, there exists $(n_k)_k \to \infty$ such that $d(F^{n_k}(\tilde{a}), \tilde{a}) \to 0$ when $k \to \infty$. This means that (1.5.2) is satisfied and, hence, the proof of Theorem 1.5.1 is complete.

1.5.2 Exercises

1.5.1. Show, by means of examples, that the conclusion of Theorem 1.5.1 is generally false if the transformations f_i do not commute with each other.

1.5.2. Let $\mathcal{G}$ be the abelian group generated by commuting homeomorphisms $f_1, \dots, f_q : M \to M$ on a compact metric space. Prove that there exists some minimal element $X \subset M$ for the inclusion relation in the family of non-empty, closed, $\mathcal{G}$-invariant subsets of M.

1.5.3. Show that if $\varphi : M \to \mathbb{R}$ is an upper semi-continuous function on a compact metric space then φ attains its maximum, that is, there exists $p \in M$ such that $\varphi(p) \geq \varphi(x)$ for every $x \in M$.

1.5.4. Show that if $\varphi : M \to \mathbb{R}$ is an (upper or lower) semi-continuous function on a compact metric space then the set of continuity points of φ contains a countable intersection of open and dense subsets of M. In particular, the set of continuity points is dense in M.

1.5.5. Let $f : M \to M$ be a measurable transformation preserving a finite measure μ. Given $k \geq 1$ and a positive measure set $A \subset M$, show that for almost every $x \in A$ there exists $n \geq 1$ such that $f^{jn}(x) \in A$ for every $1 \leq j \leq k$.

1.5.6. Let $f_1, \dots, f_q : M \to M$ be commuting homeomorphisms on a compact metric space. A point $x \in M$ is called *non-wandering* if for every neighborhood U of x there exist $n_1, \dots, n_q \geq 1$ such that $f_1^{n_1} \cdots f_q^{n_q}(U)$ intersects U. The *non-wandering set* is the set $\Omega(f_1, \dots, f_q)$ of all non-wandering points. Prove that $\Omega(f_1, \dots, f_q)$ is non-empty and compact.

Bibliography

[Aar97] J. Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.

[AB] A. Avila and J. Bochi. Proof of the subadditive ergodic theorem. Preprint www.mat.puc-rio.br/∼jairo/.

[AF07] A. Avila and G. Forni. Weak mixing for interval exchange transformations and translation flows. *Ann. of Math.*, 165:637–664, 2007.

[AKM65] R. Adler, A. Konheim, and M. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.*, 114:309–319, 1965.

[AKN06] V. Arnold, V. Kozlov, and A. Neishtadt. *Mathematical aspects of classical and celestial mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, third edition, 2006. [Dynamical systems. III], Translated from the Russian original by E. Khukhro.

[Ano67] D. V. Anosov. Geodesic flows on closed Riemannian manifolds of negative curvature. *Proc. Steklov Math. Inst.*, 90:1–235, 1967.

[Arn78] V. I. Arnold. *Mathematical methods of classical mechanics*. Springer Verlag, 1978.

[AS67] D. V. Anosov and Ya. G. Sinai. Certain smooth ergodic systems. *Russian Math. Surveys*, 22:103–167, 1967.

[Bal00] V. Baladi. *Positive transfer operators and decay of correlations*. World Scientific Publishing Co. Inc., 2000.

[BDV05] C. Bonatti, L. J. Díaz, and M. Viana. *Dynamics beyond uniform hyperbolicity*, volume 102 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, 2005.

[Bil68] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[Bil71] P. Billingsley. *Weak convergence of measures: Applications in probability*. Society for Industrial and Applied Mathematics, 1971. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 5.

[Bir13] G. D. Birkhoff. Proof of Poincaré's last Geometric Theorem. *Trans. Amer. Math. Soc.*, 14:14–22, 1913.

[Bir67] G. Birkhoff. *Lattice theory*, volume 25. A.M.S. Colloq. Publ., 1967.

[BK83] M. Brin and A. Katok. On local entropy. In *Geometric dynamics (Rio de Janeiro, 1981)*, volume 1007 of *Lecture Notes in Math.*, pages 30–38. Springer-Verlag, 1983.

[BLY] D. Burguet, G. Liao, and J. Yang. Asymptotic h-expansiveness rate of C^∞ maps. www.arxiv.org:1404.1771.

[Bos86] J.-B. Bost. Tores invariants des systèmes hamiltoniens. *Astérisque*, 133–134:113–157, 1986.

[Bos93] M. Boshernitzan. Quantitative recurrence results. *Invent. Math.*, 113(3):617–631, 1993.

[Bow71] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.*, 153:401–414, 1971.

[Bow72] R. Bowen. Entropy expansive maps. *Trans. Am. Math. Soc.*, 164:323–331, 1972.

[Bow75a] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[Bow75b] R. Bowen. A horseshoe with positive measure. *Invent. Math.*, 29:203–204, 1975.

[Bow78] R. Bowen. Entropy and the fundamental group. In *The Structure of Attractors in Dynamical Systems*, volume 668 of *Lecture Notes in Math.*, pages 21–29. Springer-Verlag, 1978.

[BS00] L. Barreira and J. Schmeling. Sets of "non-typical" points have full topological entropy and full Hausdorff dimension. *Israel J. Math.*, 116:29–70, 2000.

[Buz97] J. Buzzi. Intrinsic ergodicity for smooth interval maps. *Isreal J. Math*, 100:125–161, 1997.

[Car70] H. Cartan. *Differential forms.* Hermann, 1970.

[Cas04] A. A. Castro. *Teoria da medida.* Projeto Euclides. IMPA, 2004.

[Cla72] J. Clark. *A Kolmogorov shift with no roots.* ProQuest LLC, Ann Arbor, MI, 1972. Thesis (Ph.D.)–Stanford University.

[dC79] M. do Carmo. *Geometria riemanniana*, volume 10 of *Projeto Euclides.* Instituto de Matemática Pura e Aplicada, 1979.

[Dei85] K. Deimling. *Nonlinear functional analysis.* Springer Verlag, 1985.

[Din70] E. Dinaburg. A correlation between topological entropy and metric entropy. *Dokl. Akad. Nauk SSSR*, 190:19–22, 1970.

[Din71] E. Dinaburg. A connection between various entropy characterizations of dynamical systems. *Izv. Akad. Nauk SSSR Ser. Mat.*, 35:324–366, 1971.

[dlL93] R. de la Llave. Introduction to K.A.M. theory. In *Computational physics (Almuñécar, 1992)*, pages 73–105. World Sci. Publ., 1993.

[DS57] N. Dunford and J. Schwarz. *Linear operators I: General theory.* Wiley & Sons, 1957.

[DS63] N. Dunford and J. Schwarz. *Linear operators II: Spectral theory.* Wiley & Sons, 1963.

[Dug66] J. Dugundji. *Topology.* Allyn and Bacon Inc., 1966.

[Edw79] R. E. Edwards. *Fourier series. A modern introduction. Vol. 1*, volume 64 of *Graduate Texts in Mathematics.* Springer-Verlag, second edition, 1979.

[ET36] P. Erdös and P. Turán. On some sequences of integers. *J. London. Math. Soc.*, 11:261–264, 1936.

[Fal90] K. Falconer. *Fractal geometry.* John Wiley & Sons Ltd., 1990. Mathematical foundations and applications.

[Fer02] R. Fernandez. *Medida e integração.* Projeto Euclides. IMPA, 2002.

[FFT09] S. Ferenczi, A. Fisher, and M. Talet. Minimality and unique ergodicity for adic transformations. *J. Anal. Math.*, 109:1–31, 2009.

[FO70] N. Friedman and D. Ornstein. On isomorphism of weak Bernoulli transformations. *Advances in Math.*, 5:365–394, 1970.

[Fri69] N. Friedman. *Introduction to ergodic theory.* Van Nostrand, 1969.

[Fur61] H. Furstenberg. Strict ergodicity and transformation of the torus. *Amer. J. Math.*, 83:573–601, 1961.

[Fur77] H. Furstenberg. Ergodic behavior and a theorem of Szemerédi on arithmetic progressions. *J. d'Analyse Math.*, 31:204–256, 1977.

[Fur81] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton Univertsity Press, 1981.

[Goo71a] T. Goodman. Relating topological entropy and measure entropy. *Bull. London Math. Soc.*, 3:176–180, 1971.

[Goo71b] G. Goodwin. Optimal input signals for nonlinear-system identification. *Proc. Inst. Elec. Engrs.*, 118:922–926, 1971.

[GT08] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math.*, 167:481–547, 2008.

[Gur61] B. M. Gurevič. The entropy of horocycle flows. *Dokl. Akad. Nauk SSSR*, 136:768–770, 1961.

[Hal50] P. Halmos. *Measure Theory.* D. Van Nostrand Company, 1950.

[Hal51] P. Halmos. *Introduction to Hilbert Space and the theory of Spectral Multiplicity.* Chelsea Publishing Company, New York, 1951.

[Hay] N. Haydn. Multiple measures of maximal entropy and equilibrium states for one-simensional subshifts. Preprint Penn State University.

[Hir94] M. Hirsch. *Differential topology*, volume 33 of *Graduate Texts in Mathematics.* Springer-Verlag, 1994. Corrected reprint of the 1976 original.

[Hof77] F. Hofbauer. Examples for the nonuniqueness of the equilibrium state. *Trans. Amer. Math. Soc.*, 228:223–241, 1977.

[Hop39] E. F. Hopf. Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. *Ber. Verh. Sächs. Akad. Wiss. Leipzig*, 91:261–304, 1939.

[HvN42] P. Halmos and J. von Neumann. Operator methods in classical mechanics. II. *Ann. of Math.*, 43:332–350, 1942.

[Jac60] K. Jacobs. *Neuere Methoden und Ergebnisse der Ergodentheorie.* Ergebnisse der Mathematik und ihrer Grenzgebiete. N. F., Heft 29. Springer-Verlag, 1960.

[Jac63] K. Jacobs. *Lecture notes on ergodic theory, 1962/63. Parts I, II.* Matematisk Institut, Aarhus Universitet, Aarhus, 1963.

[Kal82] S. Kalikow. T, T^{-1} transformation is not loosely Bernoulli. *Ann. of Math.*, 115:393–409, 1982.

[Kat71] Yi. Katznelson. Ergodic automorphisms of T^n are Bernoulli shifts. *Israel J. Math.*, 10:186–195, 1971.

[Kat80] A. Katok. Lyapunov exponents, entropy and periodic points of diffeomorphisms. *Publ. Math. IHES*, 51:137–173, 1980.

[Kea75] M. Keane. Interval exchange transformations. *Math. Zeit.*, 141:25–31, 1975.

[KM10] S. Kalikow and R. McCutcheon. *An outline of ergodic theory*, volume 122 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 2010.

[Kok35] J. F. Koksma. Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins. *Compositio Math.*, 2:250–258, 1935.

[KR80] M. Keane and G. Rauzy. Stricte ergodicité des échanges d'intervalles. *Math. Zeit.*, 174:203–212, 1980.

[Kri70] W. Krieger. On entropy and generators of measure-preserving transformations. *Trans. Amer. Math. Soc.*, 149:453–464, 1970.

[Kri75] W. Krieger. On the uniqueness of the equilibrium state. *Math. Systems Theory*, 8:97–104, 1974/75.

[KSS91] A. Krámli, N. Simányi, and D. Szász. The K-property of three billiard balls. *Ann. of Math.*, 133:37–72, 1991.

[KSS92] A. Krámli, N. Simányi, and D. Szász. The K-property of four billiard balls. *Comm. Math. Phys.*, 144:107–148, 1992.

[KW82] Y. Katznelson and B. Weiss. A simple proof of some ergodic theorems. *Israel J. Math.*, 42:291–296, 1982.

[Lan73] O. Lanford. Entropy and equilibrium states in classical statistical mechanics. In *Statistical mechanics and mathematical problems*, volume 20 of *Lecture Notes in Physics*, page 1Â–113. Springer Verlag, 1973.

[Led84] F. Ledrappier. Propriétés ergodiques des mesures de Sinaï. *Publ. Math. I.H.E.S.*, 59:163–188, 1984.

[Lin77] D. Lind. The structure of skew products with ergodic group actions. *Israel J. Math.*, 28:205–248, 1977.

[LS82] F. Ledrappier and J.-M. Strelcyn. A proof of the estimation from below in pesin's entropy formula. *Ergod. Th & Dynam. Sys*, 2:203–219, 1982.

[LVY13] G. Liao, M. Viana, and J. Yang. The entropy conjecture for diffeomorphisms away from tangencies. *J. Eur. Math. Soc. (JEMS)*, 15(6):2043–2060, 2013.

[LY85a] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. *Ann. of Math.*, 122:509–539, 1985.

[LY85b] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math.*, 122:540–574, 1985.

[Man75] A. Manning. Topological entropy and the first homology group. In *Dynamical Systems, Warwick, 1974*, volume 468 of *Lecture Notes in Math.*, pages 185–190. Springer-Verlag, 1975.

[Mañ85] R. Mañé. Hyperbolicity, sinks and measure in one-dimensional dynamics. *Comm. Math. Phys.*, 100:495–524, 1985.

[Mañ87] R. Mañé. *Ergodic theory and differentiable dynamics.* Springer Verlag, 1987.

[Mas82] H. Masur. Interval exchange transformations and measured foliations. *Ann. of Math*, 115:169–200, 1982.

[Mey00] C. Meyer. *Matrix analysis and applied linear algebra.* Society for Industrial and Applied Mathematics (SIAM), 2000.

[Mis73] M. Misiurewicz. Diffeomorphim without any measure of maximal entropy. *Bull. Acad. Pol. Sci.*, 21:903–910, 1973. Series sci. math, astr. et phys.

[Mis76] M. Misiurewicz. A short proof of the variational principle for a Z_+^N action on a compact space. *Asterisque*, 40:147–187, 1976.

[MP77a] M. Misiurewicz and F. Przytycki. Entropy conjecture for tori. *Bull. Pol. Acad. Sci. Math.*, 25:575–578, 1977.

[MP77b] M. Misiurewicz and F. Przytycki. Topological entropy and degree of smooth mappings. *Bull. Pol. Acad. Sci. Math.*, 25:573–574, 1977.

[MP08] W. Marzantowicz and F. Przytycki. Estimates of the topological entropy from below for continuous self-maps on some compact manifolds. *Discrete Contin. Dyn. Syst. Ser.*, 21:501–512, 2008.

[MT78] G. Miles and R. Thomas. Generalized torus automorphisms are Bernoullian. *Advances in Math. Supplementary Studies*, 2:231–249, 1978.

[New88] S. Newhouse. Entropy and volume. *Ergodic Theory Dynam. Systems*, 8*(Charles Conley Memorial Issue):283–299, 1988.

[New90] S. Newhouse. Continuity properties of entropy. *Annals of Math.*, 129:215–235, 1990. Errata in Annals of Math. 131:409–410, 1990.

[NP66] D. Newton and W. Parry. On a factor automorphism of a normal dynamical system. *Ann. Math. Statist*, 37:1528–1533, 1966.

[NR97] A. Nogueira and D. Rudolph. Topological weak-mixing of interval exchange maps. *Ergod. Th. & Dynam. Sys.*, 17:1183–1209, 1997.

[Orn60] D. Ornstein. On invariant measures. *Bull. Amer. Math. Soc.*, 66:297–300, 1960.

[Orn70] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352 (1970), 1970.

[Orn72] D. Ornstein. On the root problem in ergodic theory. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 347–356. Univ. California Press, 1972.

[Orn74] D. Ornstein. *Ergodic theory, randomness, and dynamical systems*. Yale University Press, 1974. James K. Whittemore Lectures in Mathematics given at Yale University, Yale Mathematical Monographs, No. 5.

[OS73] D. Ornstein and P. Shields. An uncountable family of K-automorphisms. *Advances in Math.*, 10:63–88, 1973.

[OU41] J. C. Oxtoby and S. M. Ulam. Measure-preserving homeomorphisms and metrical transitivity. *Ann. of Math.*, 42:874–920, 1941.

[Par53] O. S. Parasyuk. Flows of horocycles on surfaces of constant negative curvature. *Uspehi Matem. Nauk (N.S.)*, 8:125–126, 1953.

[Pes77] Ya. B. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Math. Surveys*, 324:55–114, 1977.

[Pes97] Ya. Pesin. *Dimension theory in dynamical systems*. University of Chicago Press, 1997. Contemporary views and applications.

[Pet83] K. Petersen. *Ergodic theory*. Cambridge Univ. Press, 1983.

[Phe93] R. Phelps. *Convex functions, monotone operators and differentiability*, volume 1364 of *Lecture Notes in Mathematics*. Springer-Verlag, second edition, 1993.

[Pin60] M. S. Pinsker. *Informatsiya i informatsionnaya ustoichivostsluchainykh velichin i protsessov*. Problemy Peredači Informacii, Vyp. 7. Izdat. Akad. Nauk SSSR, 1960.

[PT93] J. Palis and F. Takens. *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations*. Cambridge University Press, 1993.

[PU10] F. Przytycki and M. Urbański. *Conformal fractals: ergodic theory methods*, volume 371 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 2010.

[PW72a] W. Parry and P. Walters. Endomorphisms of a Lebesgue space. *Bull. Amer. Math. Soc.*, 78:272–276, 1972.

[PW72b] W. Parry and P. Walters. Errata: "Endomorphisms of a Lebesgue space". *Bull. Amer. Math. Soc.*, 78:628, 1972.

[PY98] M. Pollicott and M. Yuri. *Dynamical systems and ergodic theory*, volume 40 of *London Mathematical Society Student Texts*. Cambridge University Press, 1998.

[Qua99] A. Quas. Most expanding maps have no absolutely continuous invariant mesure. *Studia Math.*, 134:69–78, 1999.

[Que87] M. Queffélec. *Substitution dynamical systems—spectral analysis*, volume 1294 of *Lecture Notes in Mathematics*. Springer-Verlag, 1987.

[Rok61] V. A. Rokhlin. Exact endomorphisms of a Lebesgue space. *Izv. Akad. Nauk SSSR Ser. Mat.*, 25:499–530, 1961.

[Rok62] V. A. Rokhlin. On the fundamental ideas of measure theory. *A. M. S. Transl.*, 10:1–54, 1962. Transl. from Mat. Sbornik 25 (1949), 107–150. First published by the A. M. S. in 1952 as Translation Number 71.

[Rok67a] V. A. Rokhlin. Lectures on the entropy theory of measure-preserving transformations. *Russ. Math. Surveys*, 22 -5:1–52, 1967. Transl. from Uspekhi Mat. Nauk. 22 - 5 (1967), 3–56.

[Rok67b] V. A. Rokhlin. Metric properties of endomorphisms of compact commuitative groups. *Amer. Math. Soc. Transl.*, 64:244–252, 1967.

[Roy63] H. L. Royden. *Real analysis*. The Macmillan Co., 1963.

[RS61] V. A. Rokhlin and Ja. G. Sinaĭ. The structure and properties of invariant measurable partitions. *Dokl. Akad. Nauk SSSR*, 141:1038–1041, 1961.

[Rud87] W. Rudin. *Real and complex analysis*. McGraw-Hill, 1987.

[Rue73] D. Ruelle. Statistical Mechanics on a compact set with Z^ν action satisfying expansiveness and specification. *Trans. Amer. Math. Soc.*, 186:237–251, 1973.

[Rue78] D. Ruelle. An inequality for the entropy of differentiable maps. *Bull. Braz. Math. Soc.*, 9:83–87, 1978.

[Rue04] D. Ruelle. *Thermodynamic formalism*. Cambridge Mathematical Library. Cambridge University Press, second edition, 2004. The mathematical structures of equilibrium statistical mechanics.

[RY80] C. Robinson and L. S. Young. Nonabsolutely continuous foliations for an Anosov diffeomorphism. *Invent. Math.*, 61:159–176, 1980.

[SC87] Ya. Sinaĭ and Nikolay Chernov. Ergodic properties of some systems of two-dimensional disks and three-dimensional balls. *Uspekhi Mat. Nauk*, 42:153–174, 256, 1987.

[Shu69] M. Shub. Endomorphisms of compact differentiable manifolds. *Amer. Journal of Math.*, 91:129–155, 1969.

[Shu74] M. Shub. Dynamical systems, filtrations and entropy. *Bull. Amer. Math. Soc.*, 80:27–41, 1974.

[Sim02] N. Simányi. The complete hyperbolicity of cylindric billiards. *Ergodic Theory Dynam. Systems*, 22:281–302, 2002.

[Sin63] Ya. Sinaĭ. On the foundations of the ergodic hypothesis for a dynamical system of Statistical Mechanics. *Soviet. Math. Dokl.*, 4:1818–1822, 1963.

[Sin70] Ya. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspehi Mat. Nauk*, 25:141–192, 1970.

[Ste58] E. Sternberg. On the structure of local homeomorphisms of Euclidean n-space - II. *Amer. J. Math.*, 80:623–631, 1958.

[SW75] M. Shub and R. Williams. Entropy and stability. *Topology*, 14:329–338, 1975.

[SX10] R. Saghin and Z. Xia. The entropy conjecture for partially hyperbolic diffeomorphisms with 1-D center. *Topology Appl.*, 157:29–34, 2010.

[Sze75] S. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[vdW27] B. van der Waerden. Beweis eibe Baudetschen Vermutung. *Nieuw Arch. Wisk.*, 15:212–216, 1927.

[Vee82] W. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.*, 115:201–242, 1982.

[Ver99] Alberto Verjovsky. *Sistemas de Anosov*, volume 9 of *Monographs of the Institute of Mathematics and Related Sciences*. Instituto de Matemática y Ciencias Afines, IMCA, Lima, 1999.

[Via14] M. Viana. *Lectures on Lyapunov Exponents*. Cambridge University Press, 2014.

[VO14] M. Viana and K. Oliveira. *Fundamentos da Teoria Ergódica*. Coleção Fronteiras da Matemática. Sociedade Brasileira de Matemática, 2014.

[Wal73] P. Walters. Some results on the classification of non-invertible measure preserving transformations. In *Recent advances in topological dynamics (Proc. Conf. Topological Dynamics, Yale Univ., New Haven, Conn., 1972; in honor of Gustav Arnold Hedlund)*, pages 266–276. Lecture Notes in Math., Vol. 318. Springer-Verlag, 1973.

[Wal75] P. Walters. A variational principle for the pressure of continuous transformations. *Amer. J. Math.*, 97:937–971, 1975.

[Wal82] P. Walters. *An introduction to ergodic theory*. Springer Verlag, 1982.

[Wey16] H. Weyl. Uber die Gleichverteilungen von Zahlen mod Eins. *Math. Ann.*, 77:313–352, 1916.

[Yan80] K. Yano. A remark on the topological entropy of homeomorphisms. *Invent. Math.*, 59:215–220, 1980.

[Yoc92] J.-C. Yoccoz. Travaux de Herman sur les tores invariants. *Astérisque*, 206:Exp. No. 754, 4, 311–344, 1992. Séminaire Bourbaki, Vol. 1991/92.

[Yom87] Y. Yomdin. Volume growth and entropy. *Israel J. Math.*, 57:285–300, 1987.

[Yos68] K. Yosida. *Functional analysis*. Second edition. Die Grundlehren der mathematischen Wissenschaften, Band 123. Springer-Verlag, 1968.

[Yuz68] S. A. Yuzvinskii. Metric properties of endomorphisms of compact groups. *Amer. Math. Soc. Transl.*, 66:63–98, 1968.

[Zyg68] A. Zygmund. *Trigonometric series: Vols. I, II.* Second edition, reprinted with corrections and some additions. Cambridge University Press, 1968.

Index