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 $\sigma$-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 $\sigma$-algebra, that is, the smallest $\sigma$-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 $\mu$ defined on the $\sigma$-algebra of $M$, such that $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$$

for any countable family $\{A_n\}$ of pairwise disjoint measurable subsets. We call $\mu$ 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 $\mu$ into a probability $\nu$: just define

$$\nu(E) = \frac{\mu(E)}{\mu(M)}$$

for every measurable set $E \subset M$.

In general, we say that a measure $\mu$ is invariant under a transformation $f$ if

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

for every measurable set $E \subset M$. \hfill (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

$$\mu(E) = \mu(f^{-t}(E))$$

for every measurable set $E \subset M$ and $t \in \mathbb{R}$. \hfill (0.0.3)

Notice that (0.0.2)–(0.0.3) do make sense, 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, ἐργόν (ergon) = work and ὁδός (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

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

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 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 Luís, 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, Felipe 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, Elais C. Malheiro, Fernando Nera Lenarduzzi, Jiagang Yang, Karina Marin, Lucas Backes, Maria João Resende, Mauricio Poletti, Paulo Varandas, Ricardo Turrolla, Sina Türelli, Vanessa Ramos, Vítor Araújo and Xiao-Chuan Liu.

Rio de Janeiro and Maceió, March 31, 2015
Marcelo Viana\textsuperscript{1} and Krerley Oliveira\textsuperscript{2}

\textsuperscript{1}IMPA, Estrada D. Castorina 110, 22460-320 Rio de Janeiro, Brasil. viana@impa.br.
\textsuperscript{2}Departamento de Matemática, Universidade Federal de Alagoas, Campus A. C. Simões s/n, 57072-900 Maceió, Brasil. krerley@mat.ufal.br.
# Contents

1 Recurrence ................................. 1
   1.1 Invariant measures ............................ 2
      1.1.1 Exercises .................................. 3
   1.2 Poincaré recurrence theorem ...................... 4
      1.2.1 Measurable version ......................... 4
      1.2.2 Kac theorem ............................... 5
      1.2.3 Topological version ....................... 7
      1.2.4 Exercises ................................. 8
   1.3 Examples ..................................... 9
      1.3.1 Decimal expansion ......................... 10
      1.3.2 Gauss map .................................. 12
      1.3.3 Circle rotations ............................ 16
      1.3.4 Rotations on tori ......................... 18
      1.3.5 Conservative maps .......................... 18
      1.3.6 Conservative flows .......................... 19
      1.3.7 Exercises .................................. 21
   1.4 Induction ..................................... 22
      1.4.1 First return map ............................ 22
      1.4.2 Induced transformations ..................... 24
      1.4.3 Kakutani-Rokhlin towers .................... 27
      1.4.4 Exercises .................................. 28
   1.5 Multiple recurrence theorems .................... 29
      1.5.1 Birkhoff multiple recurrence theorem .......... 31
      1.5.2 Exercises .................................. 34

2 Existence of Invariant Measures ................. 35
   2.1 Weak* topology ................................ 36
      2.1.1 Definition and properties of the weak* topology .... 36
      2.1.2 Portmanteau theorem ....................... 37
      2.1.3 The weak* topology is metrizable ............. 39
      2.1.4 The weak* topology is compact .............. 41
      2.1.5 Theorem of Prohorov ....................... 42
      2.1.6 Exercises .................................. 43
   2.2 Proof of the existence theorem .................. 45
2.2.1 Exercises ........................................... 48
2.3 Comments in Functional Analysis ......................... 49
  2.3.1 Duality and weak topologies .......................... 49
  2.3.2 Koopman operator .................................... 50
  2.3.3 Exercises ........................................... 52
2.4 Skew-products and natural extensions ...................... 53
  2.4.1 Measures on skew-products ............................ 54
  2.4.2 Natural extensions .................................... 54
  2.4.3 Exercises ........................................... 57
2.5 Arithmetic progressions ................................... 58
  2.5.1 Theorem of van der Waerden ........................... 60
  2.5.2 Theorem of Szemerédi ................................. 61
  2.5.3 Exercises ........................................... 63

3 Ergodic Theorems ........................................... 65
  3.1 Ergodic theorem of von Neumann .......................... 66
      3.1.1 Isometries in Hilbert spaces ......................... 66
      3.1.2 Statement and proof of the theorem ................... 69
      3.1.3 Convergence in $L^2(\mu)$ .......................... 70
      3.1.4 Exercises ........................................... 71
  3.2 Birkhoff ergodic theorem ................................ 71
      3.2.1 Mean sojourn time ................................ 71
      3.2.2 Time averages ...................................... 72
      3.2.3 Theorem of von Neumann and consequences ............. 75
      3.2.4 Exercises ........................................... 77
  3.3 Subadditive ergodic theorem .............................. 79
      3.3.1 Preparing the proof ................................. 80
      3.3.2 Key lemma .......................................... 81
      3.3.3 Estimating $\varphi_-$ ............................... 83
      3.3.4 Bounding $\varphi_+$ ................................ 84
      3.3.5 Lyapunov exponents ................................ 85
      3.3.6 Exercises ........................................... 87
  3.4 Discrete time and continuous time ......................... 88
      3.4.1 Suspension flows .................................... 89
      3.4.2 Poincaré maps ...................................... 91
      3.4.3 Exercises ........................................... 93

4 Ergodicity .................................................. 97
  4.1 Ergodic systems ........................................... 98
      4.1.1 Invariant sets and functions ......................... 98
      4.1.2 Spectral characterization ............................ 100
      4.1.3 Exercises ........................................... 103
  4.2 Examples ............................................... 104
      4.2.1 Rotations on tori ................................... 104
      4.2.2 Decimal expansion .................................. 106
      4.2.3 Bernoulli shifts .................................... 108
## CONTENTS

4.2.4 Gauss map .................................. 111
4.2.5 Linear endomorphisms of the torus ............... 114
4.2.6 Hopf argument ................................ 116
4.2.7 Exercises .................................. 119

4.3 Properties of ergodic measures ..................... 121
4.3.1 Exercises .................................. 123

4.4 Comments in Conservative Dynamics ................. 124
4.4.1 Hamiltonian systems .......................... 125
4.4.2 Kolmogorov-Arnold-Moser theory ................ 127
4.4.3 Elliptic periodic points ....................... 131
4.4.4 Geodesic flows ................................ 135
4.4.5 Anosov systems .............................. 136
4.4.6 Billiards ................................... 138
4.4.7 Exercises .................................. 144

5 Ergodic decomposition .......................... 147
5.1 Ergodic decomposition theorem ..................... 147
5.1.1 Statement of the theorem ..................... 148
5.1.2 Disintegration of a measure ................... 149
5.1.3 Measurable partitions ....................... 151
5.1.4 Proof of the ergodic decomposition theorem .... 152
5.1.5 Exercises .................................. 154
5.2 Rokhlin disintegration theorem .................... 155
5.2.1 Conditional expectations ..................... 155
5.2.2 Criterion for $\sigma$-additivity ............... 157
5.2.3 Construction of conditional measures ........... 160
5.2.4 Exercises .................................. 161

6 Unique Ergodicity .............................. 163
6.1 Unique ergodicity ............................... 163
6.1.1 Exercises .................................. 165
6.2 Minimality ................................... 165
6.2.1 Exercises .................................. 167
6.3 Haar measure ................................... 168
6.3.1 Rotations on tori ............................ 168
6.3.2 Topological groups and Lie groups ............... 169
6.3.3 Translations on compact metrizable groups .... 173
6.3.4 Odemeters .................................. 176
6.3.5 Exercises .................................. 178
6.4 Theorem of Weyl ................................ 179
6.4.1 Ergodicity .................................. 180
6.4.2 Unique ergodicity ............................ 182
6.4.3 Proof of the theorem of Weyl ................... 184
6.4.4 Exercises .................................. 186
## 7 Correlations

7.1 Mixing systems ........................................... 188
  7.1.1 Properties ............................................. 188
  7.1.2 Weak mixing .......................................... 191
  7.1.3 Spectral characterization ............................ 193
  7.1.4 Exercises ............................................. 195

7.2 Markov shifts ............................................. 196
  7.2.1 Ergodicity ............................................. 201
  7.2.2 Mixing ............................................... 203
  7.2.3 Exercises ............................................. 205

7.3 Interval exchanges ...................................... 207
  7.3.1 Minimality and ergodicity ........................... 209
  7.3.2 Mixing ............................................... 211
  7.3.3 Exercises ............................................. 214

7.4 Decay of correlations .................................. 214
  7.4.1 Exercises ............................................. 219

## 8 Equivalent Systems

8.1 Ergodic equivalence .................................... 222
  8.1.1 Exercises ............................................. 224

8.2 Spectral equivalence ................................... 224
  8.2.1 Invariants of spectral equivalence .................. 225
  8.2.2 Eigenvalues and weak mixing ....................... 226
  8.2.3 Exercises ............................................. 228

8.3 Discrete spectrum ....................................... 230
  8.3.1 Exercises ............................................. 233

8.4 Lebesgue spectrum ..................................... 233
  8.4.1 Examples and properties ............................ 233
  8.4.2 The invertible case .................................. 237
  8.4.3 Exercises ............................................. 240

8.5 Lebesgue spaces and ergodic isomorphism .............. 241
  8.5.1 Ergodic isomorphism ................................ 241
  8.5.2 Lebesgue spaces ..................................... 243
  8.5.3 Exercises ............................................. 248

## 9 Entropy

9.1 Definition of entropy ................................... 250
  9.1.1 Entropy in Information Theory ...................... 250
  9.1.2 Entropy of a partition ............................. 251
  9.1.3 Entropy of a dynamical system ...................... 256
  9.1.4 Exercises ............................................. 260

9.2 Theorem of Kolmogorov-Sinai .......................... 261
  9.2.1 Generating partitions ............................... 263
  9.2.2 Semi-continuity of the entropy ..................... 265
  9.2.3 Expansive transformations ......................... 267
  9.2.4 Exercises ............................................. 268
9.3 Local entropy ........................................... 269
  9.3.1 Proof of the Shannon-McMillan-Breiman theorem .... 270
  9.3.2 Exercises ........................................... 274
9.4 Examples ................................................ 274
  9.4.1 Markov shifts ...................................... 274
  9.4.2 Gauss map .......................................... 275
  9.4.3 Linear endomorphisms of the torus .................. 277
  9.4.4 Differentiable maps ................................ 279
  9.4.5 Exercises .......................................... 281
9.5 Entropy and equivalence .................................. 281
  9.5.1 Bernoulli automorphisms ............................ 282
  9.5.2 Systems with entropy zero ........................... 283
  9.5.3 Kolmogorov systems ................................. 286
  9.5.4 Exact systems ...................................... 291
  9.5.5 Exercises .......................................... 291
9.6 Entropy and ergodic decomposition ........................ 292
  9.6.1 Affine property .................................... 294
  9.6.2 Proof of the Jacobs theorem ......................... 296
  9.6.3 Exercises .......................................... 299
9.7 Jacobians and the Rokhlin formula ........................ 300
  9.7.1 Exercises .......................................... 305
10 Variational principle ...................................... 307
  10.1 Topological entropy .................................. 308
    10.1.1 Definition via open covers ....................... 308
    10.1.2 Generating sets and separated sets ............... 310
    10.1.3 Calculation and properties ....................... 315
    10.1.4 Exercises ...................................... 318
  10.2 Examples ............................................. 319
    10.2.1 Expansive maps .................................. 319
    10.2.2 Shifts of finite type ............................ 321
    10.2.3 Topological entropy of flows ..................... 324
    10.2.4 Differentiable maps .............................. 326
    10.2.5 Linear endomorphisms of the torus ............... 328
    10.2.6 Exercises ...................................... 330
  10.3 Pressure ............................................. 332
    10.3.1 Definition via open covers ....................... 332
    10.3.2 Generating sets and separated sets ............... 334
    10.3.3 Properties ...................................... 336
    10.3.4 Comments in Statistical Mechanics ............... 339
    10.3.5 Exercises ...................................... 343
  10.4 Variational principle .................................. 344
    10.4.1 Proof of the upper bound ......................... 346
    10.4.2 Approximating the pressure ....................... 348
    10.4.3 Exercises ...................................... 351
  10.5 Equilibrium states .................................... 352
10.5.1 Exercises ........................................... 357

11 Expanding Maps ........................................ 359
   11.1 Expanding maps on manifolds ...................... 360
      11.1.1 Distortion lemma ................................ 361
      11.1.2 Existence of ergodic measures .................. 365
      11.1.3 Uniqueness and conclusion of the proof ........ 366
      11.1.4 Exercises ........................................ 368
   11.2 Dynamics of expanding maps ....................... 369
      11.2.1 Contracting inverse branches .................... 372
      11.2.2 Shadowing and periodic points ................... 373
      11.2.3 Dynamical decomposition ......................... 377
      11.2.4 Exercises ........................................ 380
   11.3 Entropy and periodic points ...................... 381
      11.3.1 Rate growth of periodic points .................. 382
      11.3.2 Approximation by atomic measures .............. 383
      11.3.3 Exercises ........................................ 385

12 Thermodynamic Formalism ............................. 387
   12.1 Theorem of Ruelle ................................. 388
      12.1.1 Reference measure ................................ 390
      12.1.2 Distortion and the Gibbs property .............. 392
      12.1.3 Invariant density ................................ 393
      12.1.4 Construction of the equilibrium state ......... 396
      12.1.5 Pressure and eigenvalues ....................... 398
      12.1.6 Uniqueness of the equilibrium state ........... 400
      12.1.7 Exactness ........................................ 402
      12.1.8 Absolutely continuous measures ............... 403
      12.1.9 Exercises ........................................ 405
   12.2 Theorem of Livšic ................................. 406
      12.2.1 Exercises ........................................ 409
   12.3 Decay of correlations ............................. 409
      12.3.1 Projective distances ............................. 411
      12.3.2 Cones of Hölder functions ....................... 416
      12.3.3 Exponential convergence ......................... 420
      12.3.4 Exercises ........................................ 423
   12.4 Dimension of conformal repellers .................. 425
      12.4.1 Hausdorff dimension .............................. 425
      12.4.2 Conformal repellers .............................. 427
      12.4.3 Distortion and conformality ..................... 429
      12.4.4 Existence and uniqueness of \( d_0 \) ............. 431
      12.4.5 Upper bound ...................................... 434
      12.4.6 Lower bound ...................................... 435
      12.4.7 Exercises ........................................ 436
A Measure Theory, Topology and Analysis 439

A.1 Measure spaces ........................................ 439
A.1.1 Measurable spaces ..................................... 440
A.1.2 Measure spaces ......................................... 442
A.1.3 Lebesgue measure ....................................... 445
A.1.4 Measurable maps ....................................... 448
A.1.5 Exercises ................................................ 450

A.2 Integration in measure spaces ........................... 452
A.2.1 Lebesgue integral ....................................... 453
A.2.2 Convergence theorems .................................. 455
A.2.3 Product measures ....................................... 456
A.2.4 Derivation of measures ................................. 458
A.2.5 Exercises ................................................ 460

A.3 Measures in metric spaces ................................ 462
A.3.1 Regular measures ....................................... 462
A.3.2 Separable complete metric spaces ..................... 465
A.3.3 Space of continuous functions ......................... 466
A.3.4 Exercises ................................................ 468

A.4 Differentiable manifolds .................................. 469
A.4.1 Differentiable manifolds and maps ..................... 469
A.4.2 Tangent space and derivative .......................... 471
A.4.3 Cotangent space and differential forms ................ 473
A.4.4 Transversality .......................................... 474
A.4.5 Riemannian manifolds ................................ 475
A.4.6 Exercises ................................................ 477

A.5 $L^p(\mu)$ spaces .......................................... 478
A.5.1 $L^p(\mu)$ spaces with $1 \leq p < \infty$ ................. 478
A.5.2 Inner product in $L^2(\mu)$ ............................. 479
A.5.3 Space of essentially bounded functions ............... 480
A.5.4 Convexity ............................................... 480
A.5.5 Exercises ................................................ 481

A.6 Hilbert spaces ............................................. 482
A.6.1 Orthogonality .......................................... 482
A.6.2 Duality .................................................. 483
A.6.3 Exercises ................................................ 485

A.7 Spectral theorems ......................................... 486
A.7.1 Spectral measures ....................................... 486
A.7.2 Spectral representation ................................. 489
A.7.3 Exercises ................................................ 490

Hints or solutions for selected exercises 491
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 trans-
formation 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 transfor-
mation. We say that the measure \(\mu\) is invariant under \(f\) if

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

We also say that \(\mu\) is \(f\)-invariant, or that \(f\) preserves \(\mu\), 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:

\[ f^0 = \text{id} \quad \text{and} \quad f^{s+t} = f^s \circ f^t \quad \forall s, t \in \mathbb{R}. \quad (1.1.2) \]

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 \(\mu\) is invariant under a flow \((f^t)\), if it is invariant
under each one of the transformations \(f^t\), that is, if

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

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

\[ \int \phi \, d\mu = \int \phi \circ f \, d\mu \quad (1.1.4) \]

for every \(\mu\)-integrable function \(\phi : M \to \mathbb{R}\).
1.1. INVARIANT MEASURES

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

\[
\mu(B) = \int \chi_B \, d\mu \quad \text{and} \quad \mu(f^{-1}(B)) = \int \chi_{f^{-1}(B)} \, d\mu = \int (\chi_B \circ f) \, d\mu.
\]

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 \( \phi \) and such that \( |s_n| \leq |\phi| \) 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 \( \mu \) 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 \( \delta_p \) is invariant under \( f \) if and only if \( p \) is a fixed point of \( f \). More generally, a probability measure \( \delta_{p,k} = k^{-1}(\delta_p + \delta_{f(p)} + \cdots + \delta_{f^{k-1}(p)}) \) 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 \( \mu \) be a measure on \( M \). Show that \( f \) preserves \( \mu \) 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 \( \mu \) then, given any \( k \geq 2 \), the iterate \( f^k \) also preserves \( \mu \). Is the converse true?

1.1.4. Suppose that \( f : M \to M \) preserves a probability measure \( \mu \). 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 \( \mu \) be a finite measure invariant under \( f \). Let \( E \subset M \) be any measurable set with \( \mu(E) > 0 \). Then, for \( \mu \)-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 \( f^{m-n}(y) = f^n(x) \in E_0 \). 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 \( \mu \) is invariant, we also have that \( \mu(f^{-n}(E_0)) = \mu(E_0) \) for all \( n \geq 1 \). It follows that

\[
\mu\left( \bigcup_{n=1}^{\infty} f^{-n}(E_0) \right) = \sum_{n=1}^{\infty} \mu(f^{-n}(E_0)) = \sum_{n=1}^{\infty} \mu(E_0).
\]

The expression on the left-hand side is finite, since the measure \( \mu \) 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 \( \mu \) is invariant, it follows that

\[
\mu(F) \leq \mu\left( \bigcup_{k=0}^{\infty} f^{-k}(E_0) \right) \leq \sum_{k=0}^{\infty} \mu(f^{-k}(E_0)) = \sum_{k=0}^{\infty} \mu(E_0) = 0.
\]
1.2. POINCARÉ RECURRENCE THEOREM

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

Theorem 1.2.1 implies an analogous result for continuous time systems: if \( \mu \) is a finite invariant measure of a flow \((f^t)\), then for every measurable set \( E \subset M \) with positive measure and for \( \mu \)-almost every \( x \in E \), there exist times \( t_j \to +\infty \) such that \( f^{t_j}(x) \in E \). Indeed, if \( \mu \) 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 \( \mu \) 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

\[
\rho_E(x) = \min\{n \geq 1 : f^n(x) \in E\}
\]  

(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:

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

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, \( \mu \) be a finite invariant measure and \( E \subset M \) be a positive measure set. Then the function \( \rho_E \) is integrable and

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

**Proof.** For each \( n \geq 1 \), define

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

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

\[
E_n = \{ x \in E : \rho_E(x) = n \},
\]
CHAPTER 1. RECURRENCE

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, $\rho_E$ 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,

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

Now observe that

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

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 $y \in E_n^{n+1}$. This proves the equality (1.2.3). So, given that $\mu$ 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

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

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

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

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

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

as we wanted to prove.  

In some cases, for example when the system $(f, \mu)$ 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 Kac theorem means that

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

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 $\sigma$-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 $(f^t)_{t \in \mathbb{R}}$ 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 $\mu$ be a finite measure on $M$ invariant under $f$. Then, $\mu$-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 $x \notin U_k$. 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 $\mu$ 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 $X_\alpha$ 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 $\mu$ 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 $\mu$ 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, $\mu$-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 $\mu$ is a $\sigma$-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, $\mu$-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, $\mu$ 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 $D$:

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

for $\mu$-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 $\mu$. 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 $V_A = \{ n_1, n_2, \ldots \}$ is a syndetic, that is, that there exists $C > 0$ such that $n_{i+1} - n_i \leq C$ for every $i$.

1. Show that for any increasing sequence $k_1 < k_2 < \cdots$ there exist $j > i \geq 1$ 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 $\ell$. Show that $V_A$ intersects $S(\ell)$ for every $\ell$.

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

$$\liminf \frac{n}{n} |f^n(x) - x| \leq 1.$$  

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

1.2.7. Define $f : [0, 1] \to [0, 1]$ by $f(x) = (x + \omega) - [x + \omega]$, where $\omega$ 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 $\omega$ and $\omega - \sqrt{5}$, prove that $\liminf_n n^2 |\omega - q_n| \geq 1/\sqrt{5}$. [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:

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

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](image)

We claim that the Lebesgue measure \(\mu\) on the interval is invariant under the transformation \(f\), that is, it satisfies

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

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 \(\sigma\)-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 \(\mu\) be a finite measure on \(M\). Suppose that there exists some algebra \(A\) of measurable subsets of \(M\) such that \(A\) generates the \(\sigma\)-algebra \(B\) of \(M\) and \(\mu(E) = \mu(f^{-1}(E))\) for every \(E \in A\). Then the latter remains true for every set \(E \in B\), that is, the measure \(\mu\) is invariant under \(f\).

**Proof.** We start by proving that \(C = \{ E \in B : \mu(E) = \mu(f^{-1}(E)) \}\) is a monotone class. Let \(E_1 \subset E_2 \subset \ldots\) be any increasing sequence of elements of \(C\) and
1.3. EXAMPLES

Let \( E = \bigcup_{i=1}^{\infty} E_i \). 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 \( \sigma \)-algebra \( \mathcal{B} \) generated by \( \mathcal{A} \). That is precisely what we wanted to prove. \( \square \)

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

\[
x = 0.a_0a_1a_2a_3\cdots
\]

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

\[
f(x) = 0.a_1a_2a_3\cdots.
\]

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

\[
f^n(x) = 0.a_na_{n+1}a_{n+2}\cdots \quad \text{(1.3.2)}
\]

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

\[
a_1 = \left\lfloor \frac{1}{x_0} \right\rfloor \quad \text{and} \quad x_1 = \frac{1}{x_0} - a_1.
\]

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

\[
a_2 = \left\lfloor \frac{1}{x_1} \right\rfloor \quad \text{and} \quad x_2 = \frac{1}{x_1} - a_2.
\]

Then

\[
x_1 = \frac{1}{a_1 + x_2} \quad \text{and so} \quad x_0 = \frac{1}{\frac{1}{a_1 + x_1}}.
\]

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

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

and observe that

\[
x_0 = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ldots + \frac{1}{a_n + x_n}}}}. \tag{1.3.3}
\]

It can be shown that the sequence

\[
z_n = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ldots + \frac{1}{a_n}}}} \tag{1.3.4}
\]

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

\[
x_0 = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ldots + \frac{1}{a_n + \ldots}}}}. \tag{1.3.5}
\]
which is called \textit{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 \textit{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 \textit{Gauss map} \(G : [0,1] \to [0,1]\) is defined by

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

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 \(G(x) = 1/x - k\).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{gauss_map.png}
\caption{Gauss map}
\end{figure}

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

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

(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

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

(1.3.7)

In particular, $\mu$ is indeed equivalent to the Lebesgue measure $m$.

**Proposition 1.3.2.** The measure $\mu$ is invariant under $G$. Moreover, if we choose $c = 1/\log 2$ then $\mu$ 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 $\bigcup_k I_k$ has full Lebesgue measure in $[0,1]$ and
2. the restriction $f_k = f | I_k$ 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

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

(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,

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

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

$$
\int_0^1 \phi(f(x))\rho(x) dx = \sum_{k=1}^{\infty} \int_{I_k} \phi(f(x))\rho(x) dx
$$

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

(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 $\phi = X_E$, this means that $\mu(f^{-1}(E)) = \mu(E)$ for every measurable set $E \subset [0,1]$. In other words, $\mu$ 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,

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

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 $\mu$ 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 $\mu$ 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 $n \in N$. 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:

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

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

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

(1.3.11)

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

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

Note that $R_\theta$ corresponds, via the identification (1.3.11), to the transformation $z \mapsto e^{2\pi i \theta}z$ 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 $\theta$ and $\tau$. In particular, every $R_\theta$ 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 $\pi: \mathbb{R} \rightarrow S^1$ 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 $\mu$ on the circle to be given by

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

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

It is clear from the definition that $\mu$ is a probability. Moreover, $\mu$ is invariant under every rotation $R_\theta$ (according to Exercise 1.3.8, it is the only probability measure with this property). This can be shown as follows. By definition,
1.3. EXAMPLES

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

\[
m((\pi^{-1}(E) - \theta) \cap [0, 1)) = m(\pi^{-1}(E) \cap [\theta, \theta + 1))
= m(\pi^{-1}(E) \cap [\theta, k+1)) + m(\pi^{-1}(E) \cap [k+1, \theta + 1)).
\]

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

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

Combining these two relations we find that

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

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 \(\theta\). If \(\theta\) is rational, say \(\theta = p/q\) with \(p \in \mathbb{Z}\) and \(q \in \mathbb{N}\), then

\[
R_\theta^q([x]) = [x + q\theta] = [x]
\]

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 \(\theta\) is irrational then \(O([x]) = \{R_\theta^n([x]) : n \in \mathbb{N}\}\) 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 \(\theta\) 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 \(0 \leq r - la < a\). This implies that

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

As \(m = lq\) and \(n = -lq\) are integers and \(\varepsilon\) 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 \(\varepsilon\). Since \(x, y\) and \(\varepsilon\) are arbitrary, this shows that every orbit \(O([x])\) is dense in \(S^1\). \(\square\)

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

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

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 $R^d_\theta : \mathbb{T}^d \to \mathbb{T}^d$, $R^d_\theta([x]) = [x] + [\theta]$ the rotation by $\theta$ (sometimes, $R^d_\theta$ is also called the translation by $\theta$). 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 $\mu$ on the $d$-torus, through the following formula:

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

This measure $\mu$ is invariant under $R^d_\theta$ for every $\theta$. 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

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

Otherwise, we say that $\theta$ is rationally dependent. One can show that $\theta$ is rationally independent if and only if the rotation $R^d_\theta$ is minimal, meaning that the orbit $O([x]) = \{R^d_\theta([x]) : n \in \mathbb{N}\}$ 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 $\text{vol}$ 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 $B \subset M$,

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

One can easily deduce the following consequence:
1.3. Examples

Lemma 1.3.5. A $C^1$ diffeomorphism $f: M \to M$ preserves the volume measure $\text{vol}$ 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

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

This means that $f$ preserves the measure $\text{vol}$ 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

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

Denote $E = f(U)$. Since $\text{vol}(U) > 0$, the previous inequality implies that $\text{vol}(E) > \text{vol}(f^{-1}(E))$. 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

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

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 \( \text{div} \ F \) of the vector field:

\[
\det Df^t(x) = \exp \left( \int_0^t \text{div} \ F(f^s(x)) \, ds \right) \quad \text{for every } x \text{ and every } t.
\]

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

\[
\text{div} \ F = \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_d}{\partial x_d}. \tag{1.3.15}
\]

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 \( \omega \), called the volume form (this remains true in the non-orientable case, except that the form \( \omega \) 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

\[
\text{vol}(B) = \int_B \rho(x_1, \ldots, x_d) \, dx_1 \cdots dx_d,
\]

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

\[
\text{div} \ F = \frac{\partial (\rho F)}{\partial x_1} + \cdots + \frac{\partial (\rho F)}{\partial x_d}
\]

(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 \((f^t)\), associated with a \(C^1\) vector field \(F\) on a Riemannian manifold preserves the volume measure on the manifold if and only if \(\text{div} F = 0\) 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 \(\text{div} F = 0\), 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, \(x = 0.3375617264\cdot\cdot\cdot\)), 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 \(f : M \rightarrow M\) be a \(C^1\) local diffeomorphism on a compact Riemannian manifold \(M\). Let \(\text{vol}\) be the volume measure on \(M\) and \(\rho : M \rightarrow [0, \infty)\) be a continuous function. Then \(f\) preserves the measure \(\mu = \rho \text{vol}\) 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 \(\mu\) 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 \(f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d\) 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 \(\theta\), the Lebesgue measure is the unique probability measure invariant under \(R_\theta\).]
CHAPTER 1. RECURRENCE

1.3.9. Suppose that \( \theta = (\theta_1, \ldots, \theta_d) \) 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 \( R_\theta \), in the sense that \( R_\theta(U) = U \) and \( R_\theta(V) = V \). Deduce that no orbit \( \mathcal{O}([x]) \) of the rotation \( R_\theta \) is dense in \( \mathbb{T}^d \).

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

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 \((p_1, \ldots, p_d, q_1, \ldots, q_d)\) the coordinate variables in \( \mathbb{R}^{2d} \). The Hamiltonian vector field associated with \( H \) is defined by

\[
F(p_1, \ldots, p_d, q_1, \ldots, q_d) = \left( \frac{\partial H}{\partial q_1}, \ldots, \frac{\partial H}{\partial q_d}, -\frac{\partial H}{\partial p_1}, \ldots, -\frac{\partial H}{\partial p_d} \right).
\]

Check that the flow defined by \( F \) preserves the volume measure.

1.4 Induction

In this section we describe a general method, based on the Poincaré recurrence theorem, to construct from a given system \((f, \mu)\) other systems, that we refer to as systems induced by \((f, \mu)\). 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 \( \mu \) 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 $\mu$-almost every $x \in E$ and so $g$ is defined on a full measure subset of $E$. We also denote by $\mu_E$ the restriction of $\mu$ to the measurable subsets $E$.

**Proposition 1.4.1.** The measure $\mu_E$ 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

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

On the other hand, since $\mu$ is $f$-invariant,

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

Analogously,

$$\mu(f^{-1}(B) \setminus E) = \mu(f^{-2}(B) \setminus f^{-1}(E))$$

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

Replacing this expression in (1.4.2), we find that

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

Repeating this argument successively, we obtain

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

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

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

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 $\mu_E$ is invariant under $g$. \hfill $\square$

**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 $\rho$ 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 $\nu$ invariant under $g : E \to E$, we may construct a certain related measure $\nu_\rho$ 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
\[ g : E \to E, \quad g(x) = f^{\rho(x)}(x), \quad (1.4.4) \]
where $\rho : E \to \mathbb{N}$ is a measurable function (it suffices that $\rho$ 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
\[ \nu_\rho(B) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(f^{-n}(B) \cap E_k), \quad (1.4.5) \]
for every measurable set $B \subset M$.

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

**Proof.** First, let us prove that $\nu_\rho$ is invariant. By the definition (1.4.5),
\[ \nu_\rho(f^{-1}(B)) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(f^{-(n+1)}(B) \cap E_k) = \sum_{n=1}^{\infty} \sum_{k \geq n} \nu(f^{-n}(B) \cap E_k). \]

We may rewrite this expression as follows:
\[ \nu_\rho(f^{-1}(B)) = \sum_{n=1}^{\infty} \sum_{k \geq n} \nu(f^{-n}(B) \cap E_k) + \sum_{k=1}^{\infty} \nu(f^{-k}(B) \cap E_k). \quad (1.4.6) \]

Concerning the last term, observe that
\[ \sum_{k=1}^{\infty} \nu(f^{-k}(B) \cap E_k) = \nu(g^{-1}(B)) = \nu(B) = \sum_{k=1}^{\infty} \nu(B \cap E_k), \]
1.4. INDUCTION

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

\[
\nu_f(B) = \sum_{n=1}^{\infty} \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \sum_{k=1}^{\infty} \nu(B \cap E_k) = \nu_f(B)
\]

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

\[
\nu_f(M) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(f^{-n}(M) \cap E_k) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(E_k) = \sum_{k=1}^{\infty} k \nu(E_k) = \int_E \rho \, dv.
\]

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 \( \nu \) is the restriction \( \mu | E \) of some invariant measure \( \mu \) of \( f \):

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

1. \( \nu_f(B) = \nu(B) = \mu(B) \) for every measurable set \( B \subset E \).
2. \( \nu_f(B) \leq \mu(B) \) 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, \( \nu_f(B) = \sum_{k>0} \nu(B \cap E_k) = \nu(B) \) as claimed in the first part of the statement.

Consider any measurable set \( B \subset M \). Then,

\[
\mu(B) = \mu(B \cap E) + \mu(B \cap E^c) = \nu(B \cap E) + \mu(B \cap E^c)
\]

\[
= \sum_{k=1}^{\infty} \nu(B \cap E_k) + \mu(B \cap E^c).
\]

(1.4.7)

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

\[
\mu(B \cap E^c) = \mu(f^{-1}(B) \cap E \cap f^{-1}(E^c)) + \mu(f^{-1}(B) \cap E^c \cap f^{-1}(E^c))
\]

\[
= \sum_{k=2}^{\infty} \nu(f^{-1}(B) \cap E_k) + \mu(f^{-1}(B) \cap E^c \cap f^{-1}(E^c)).
\]

Replacing this in (1.4.7), we find that

\[
\mu(B) = \sum_{n=0}^{1} \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \mu(f^{-1}(B) \cap \bigcap_{n=0}^{1} f^{-n}(E^c)).
\]
Repeating this argument successively, we get that

\[
\mu(B) = \sum_{n=0}^{N} \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \mu(f^{-N}(B) \cap \bigcap_{k=0}^{N} f^{-n}(E^c)) \\
\geq \sum_{n=0}^{N} \sum_{k>n} \nu(f^{-n}(B) \cap E_k) \quad \text{for every } N \geq 1.
\]

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

We also have from the Kac 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 \( a(1 + ad) = 1 \). Then define \( f : [0, 1] \to [0, 1] \) as follows:

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

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.

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

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

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 \( \nu \) equivalent to the Lebesgue measure on \((0,1]\). In fact, the density (Radon-Nikodym derivative) of \( \nu \) with respect to the Lebesgue measure is bounded from zero and infinity. Then, the \( f \)-invariant measure \( \nu_\rho \) 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 \( \nu \) 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 \( \nu_\rho \) 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 \( \rho \), and the restriction of \( \nu_\rho \) to \( E \) coincides with \( \nu \).

This transformation \( f \) is called the Kakutani-Rokhlin tower of \( g \) with time \( \rho \). The measure \( \nu_\rho \) is finite if and only if \( \rho \) is integrable with respect to \( \nu \). They are constructed as follows. Begin by defining

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

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 \( \nu_\rho \) is defined by

\[
\nu_\rho | (E_k \times \{n\}) = \nu | E_k
\]
for every $0 \leq n < k$. It is clear that the restriction of $\nu_\rho$ to the ground floor coincides with $\nu$. Moreover, $\nu_\rho$ is invariant under $f$ and

$$\nu_\rho(M) = \sum_{h=1}^{\infty} h \nu(E_h) = \int_E \rho \, dv.$$  

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 $\tau(x) = \min\{k \geq 0 : f^k(x) \in (1/2, 1)\}$ 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 $\nu_\rho$ and the sequence $(a_n)_n$ defined in Example 1.4.5. Check that $\nu_\rho$ is always $\sigma$-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

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

Deduce that the $g$-invariant measure $\nu_\rho$ 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 $\Sigma = \{1, \ldots, d\}^\mathbb{Z}$ 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 $\mu$ 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
1.5. MULTIPLE RECURRENCE THEOREMS

Let \( n \geq 1 \) 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 \( \varepsilon \). 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, \( H = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} H_j \). 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

\[
\lim_{n} \frac{1}{n} \left\{ 1 \leq j \leq n : x \in H_j \right\} \geq \theta > 0 \quad \Rightarrow \quad \lim_{k} \inf \frac{1}{k} \sum_{i=0}^{k-1} \tau(T^i(x)) \leq \frac{1}{\theta}.
\]

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

\[
\lim_{n} \frac{1}{n} \left\{ 1 \leq j \leq n : x \in H_j \right\} \geq \theta > 0 \quad \text{for} \quad \mu\text{-almost every} \quad x \in M.
\]

Show that \( \tau_{n+1}(x)/\tau_n(x) \to 1 \) for \( \mu \)-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

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

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

\[
\lim_{k} f_i^{n_k}(a) = a \quad \text{for every} \quad i = 1, \ldots, q. \tag{1.5.1}
\]

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 \(\mu\). Then, given any set \(E \subset M\) with positive measure, there exists \(n \geq 1\) such that

\[
\mu\left(E \cap f_1^{-n}(E) \cap \cdots \cap f_q^{-n}(E)\right) > 0.
\]

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

\[
\mu\left(E \cap f_1^{-(m+n)}(E) \cap \cdots \cap f_q^{-(m+n)}(E)\right) > 0.
\]

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 \(\mu\)-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 \(\mu\) 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 \(\hat{U}_k \subset U_k\) with zero measure such that every point in \(U_k \setminus \hat{U}_k\) has infinitely many simultaneous returns to \(U_k\). Then \(\hat{U} = \cup_k \hat{U}_k\) has measure zero and every point in its complement is simultaneously recurrent, in the sense of Theorem 1.5.1.
1.5. MULTIPLE RECURRENCE THEOREMS

1.5.1 Birkhoff multiple recurrence theorem

In this section we prove Theorem 1.5.1 in the case when the transformations
\( f_1, \ldots, f_q \) 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 \( F(x_1, \ldots, x_q) = (f_1(x_1), \ldots, f_q(x_q)) \). Denote by \( \Delta_q \) the diagonal of \( M^q \), that is, the subset of points of the form \( \tilde{x} = (x, \ldots, x) \). Theorem 1.5.1 claims, precisely, that there exist \( \tilde{a} \in \Delta_q \) and \((n_k)_k \to \infty\) such that
\[
\lim_{k} F^{n_k}(\tilde{a}) = \tilde{a}.
\] (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 \( f_1, \ldots, f_q \).

Let \( G \) be the (abelian) group generated by the homeomorphisms \( f_1, \ldots, f_q \). We say that a set \( X \subset M \) is \( G \)-invariant if \( g(X) \subset X \) for every \( g \in G \). Observing that the inverse \( g^{-1} \) is also in \( G \), we see that this implies \( g(X) = X \) for every \( g \in G \). 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, \( G \)-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 \( H \subset G \) such that
\[
\bigcup_{h \in H} h^{-1}(U) = M.
\]

**Proof.** For any \( x \in M \), the closure of the orbit \( \mathcal{G}(x) = \{g(x) : g \in G\} \) is a non-empty, closed, \( G \)-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 \( g \in G \) such that \( g(x) \in U \). This proves that \( \{g^{-1}(U) : g \in G\} \) 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
\[
d((x_1, \ldots, x_q), (y_1, \ldots, y_q)) = \max\{d(x_i, y_i) : 1 \leq i \leq q\}.
\]
Note that the map \( M \to \Delta_q, x \mapsto \tilde{x} = (x, \ldots, x) \) is a homeomorphism, and even an isometry for this choice of a distance. Every open set \( U \subset M \) corresponds to an open set \( \tilde{U} \subset \Delta_q \) through this homeomorphism. Given any \( g \in \mathcal{G} \), we denote by \( \tilde{g} : M^q \to M^q \) the homeomorphism defined by \( \tilde{g}(x_1, \ldots, x_q) = (g(x_1), \ldots, g(x_q)) \). The fact that the group \( G \) is abelian implies that \( \tilde{g} \) commutes
with $F$; note also that every $\tilde{g}$ preserves the diagonal $\Delta_q$. Then the conclusion of Lemma 1.5.3 may be rewritten in the following form:

$$\bigcup_{h \in \mathcal{H}} \tilde{h}^{-1}(\tilde{U}) = \Delta_q.$$  \hspace{1cm} (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,

$$F^{n_k}(\tilde{x}_k) = (f_1^{n_k} f_q^{-n_k}(y), \ldots, f_{q-1}^{n_k} f_q^{-n_k}(y), f_q^{n_k} f_q^{-n_k}(y))$$

$$= (g_1^{n_k}(y), \ldots, g_{q-1}^{n_k}(y), y)$$

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. \hfill $\square$

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 $\tilde{U} = \text{open ball of center } \tilde{z}$ 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 $\tilde{h}^{-1}(\tilde{U})$, $h \in \mathcal{H}$ cover $\Delta_q$. 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 $\tilde{y} \in \tilde{h}^{-1}(\tilde{U})$. Then,

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

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. \hfill $\square$

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 $\varepsilon_j$, $m_j$ and $\tilde{z}_j$, $j \geq 1$ defined by recurrence as follows. Initially, take $\varepsilon_1 = \varepsilon/2$.
In particular, for any $i < j$,

$$d(F^m(\tilde{z}_i), \tilde{z}_0) < \varepsilon_1.$$ 

This completes the proof of the lemma.

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 $d(\tilde{z}_j, \tilde{z}_{j+1}) < \varepsilon_{j+1}$ implies $d(F^{m_j}(\tilde{z}_j), \tilde{z}_{j-1}) < \varepsilon_j$.

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 $\Delta_q$ 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 $\phi$ is upper semi-continuous: given any $\varepsilon > 0$, every point $\tilde{x}$ admits some neighborhood $V$ such that $\phi(y) < \phi(\tilde{x}) + \varepsilon$ for every $y \in V$. This is an immediate consequence of the fact that $\phi$ is given by the infimum of a family of continuous functions. Then (Exercise 1.5.4), $\phi$ 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 $\phi(y) \geq \beta > 0$ for every $y \in V$. Then,

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

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

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

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

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

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 $f_1, \ldots, f_q : M \to M$ on a compact metric space. Prove that there exists some minimal element $X \subseteq M$ 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 $\varphi$ 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 $\varphi$ 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 $\mu$. 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 $f_1, \ldots, f_q : M \to M$ 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 $n_1, \ldots, n_q \geq 1$ such that $f_1^{n_1} \cdots f_q^{n_q}(U)$ intersects $U$. The non-wandering set is the set $\Omega(f_1, \ldots, f_q)$ of all non-wandering points. Prove that $\Omega(f_1, \ldots, f_q)$ is non-empty and compact.
Bibliography


BIBLIOGRAPHY


BIBLIOGRAPHY

BIBLIOGRAPHY


Index

$2^X$  family of all subsets, 440
$A \Delta B$ symmetric difference of sets, 444
$B(x,T;\varepsilon)$ dynamical ball for flows, 324
$B(x,\infty;\varepsilon)$ infinite dynamical ball, 331
$B(x,n;\varepsilon)$ dynamical ball, 269
$B(x,r)$ ball of center $x$ and radius $r$, 462
$B^\delta$ $\delta$-neighborhood of a set, 36
$C^*$ dual cone, 52
$C^0(M)$ space of continuous functions, 50, 445, 466
$C^0_+(M)$ cone of positive functions, 52
$C^0(M)$ space of Hölder functions, 422
$C^r(M,N)$ space of $C^r$ maps, 470
$C_n(\varphi,\psi)$ correlations sequence, 188
$D_l$ lower density, 59
$D_u$ upper density, 59
$D(f)$ derivative of a map, 471, 472
$E(A,P)$ conditional expectation, 157
$E^*$ dual of a Banach space, 49
$G(f,\phi)$ pressure, via generating sets, 335
$H(\alpha)$ entropy of an open cover, 308
$H_\beta(g)$ Hölder constant, 409
$H_\mu(P)$ entropy of a partition, 252
$H_\mu(P/Q)$

conditional entropy, 254
$H_{\beta,\rho}(g)$ local Hölder constant, 409
$I(U)$ set of invariant vectors, 67
$I(A)$ mean information of an alphabet, 251
$I(\alpha)$ information of a symbol, 251
$I_P$ information function of a partition, 252
$L^\infty(\mu)$ space of essentially bounded functions, 480
$L^p(\mu)$ space of $p$-integrable functions, 478
$P(f,\phi)$ pressure, 332
$P(f,\phi,\alpha)$ pressure relative to an open cover, 332
$P(x,\cdot)$ transition probability, 196
$P_{i,j}$ transition probability, 197
$R_\theta$ rotation on circle or torus, 16
$S(f,\phi)$ pressure, via separated sets, 335
$S^1$ circle, 16
$S^d$ orthogonal complement, 483
$S^d$ sphere of dimension $d$, 469
$TM$ tangent bundle, 472
$T^3M$ unit tangent bundle, 477
$T_pM$ tangent space at a point, 471
$U_f$ Koopman operator, 50, 51
$U_f^*$ dual of the Koopman operator, 51
$V(\mu,\Phi,\varepsilon)$

neighborhoods in weak* topology, 36
INDEX

$V(v, \{g_1, \ldots, g_N\}, \varepsilon)$
neighborhoods in weak topology, 49

$V^*(g, \{v_1, \ldots, v_N\}, \varepsilon)$
neighborhoods in weak* topology, 50

$V_0(\mu, A, \varepsilon)$
neighborhoods in weak* topology, 37

$V_\varepsilon(\mu, B, \varepsilon)$
neighborhoods in weak* topology, 37

$V_f(\mu, F, \varepsilon)$
neighborhoods in weak* topology, 37

$V_p(\mu, B, \varepsilon)$
neighborhoods in pointwise topology, 44

$V_u(\mu, \varepsilon)$
neighborhoods in uniform topology, 44

$\chi_B$ characteristic function of a set, 449

Diffeo$^r(M)$ space of $C^r$ diffeomorphisms, 471

Fix$(f)$ set of fixed points, 320

GL$(d, \mathbb{R})$ linear group, 79, 170, 474

O$(d, \mathbb{R})$ orthogonal group, 170

SL$(d, \mathbb{R})$ special linear group, 170, 475

$\Sigma_A$, $\Sigma_P$ shift of finite type, 199, 321

$\alpha \lor \beta$ sum of open covers, 308

$\alpha \prec \beta$ order relation for open covers, 308

$\alpha^n, \alpha^\pm n$ iterated sum of an open cover, 308, 315, 316

$L^1(\mu)$ space of integrable functions, 453

$M(X)$ space of measures, 50, 445

$M_1(M)$ space of probability measures, 36

$M_1(f)$ space of invariant probability measures, 121

$M_1(\mathcal{F})$ space of ergodic probability measures, 121

$\mathcal{P} \prec \mathcal{Q}$ order relation for partitions, 254

$\mathcal{P} \lor \mathcal{Q}$ sum of partitions, 252

$\mathcal{P}^n, \mathcal{P}^\pm n$ iterated sum of a partition, 256, 259

$U^r(f, \varepsilon)$ $C^r$ neighborhood of a map, 470

$\delta_\mu$ Dirac measure, 47

div $F$ divergence of a vector field, 20

degree$(f)$ degree of a map, 477

$\lambda = (\lambda_n)_n$ length vector, 208

$\lambda_{\text{max}}$ largest Lyapunov exponent, 86

$\lambda_{\text{min}}$ smallest Lyapunov exponent, 86

$\mu \perp \nu$ mutually singular measures, 459

$\nu \ll \mu$ absolutely continuous measure, 459

$\partial I$ left endpoint of an interval, 209

$\partial \mathcal{P}$ boundary of a partition, 265, 349

$\partial d$ projective space, 477

$\mu(B)$ spectral radius, 323

$\text{spec}(L)$ spectrum of a linear operator, 486

supess $\varphi$ essential supremum of a function, 480

$\text{supp} \mu$ support of a measure, 448, 486

$tanh$ hyperbolic tangent, 414

$\tau(E, x)$ mean sojourn time, 65

$\theta(g_1, g_2)$ projective distance, 412

$\check{\varphi}$ time average of a function, 73

$\tau^d$ torus of dimension $d$, 18, 469

$\varphi^+$ positive part of a function, 453

$\varphi^-$ negative part of a function, 453

$\varphi_n$ orbital sum of a function, 332, 388

$\vee_{\alpha U_\alpha}$ $\sigma$-algebra generated by a family, 286

$d$-adic interval, 498

d$(M)$ Hausdorff dimension, 425

$e(\psi, x)$ conditional expectation, 155

$f_\mu$ image of a measure, 45, 50

$f_A$ linear endomorphism of the torus, 115

$g(\phi)$
INDEX

topological entropy of flows, via generating sets, 325
\(g(f)\)
topological entropy, via generating sets, 311
\(h(f)\)
topological entropy, 309
\(h(f,\alpha)\)
entropy relative to an open cover, 309
\(h_{\mu}(f)\)
entropy of a dynamical system, 257
\(h_{\mu}(f, P)\)
entropy relative to a partition, 257
\(h_{\mu}(f, P, x)\)
entropy (local) at a point, 269
\(m_d(M)\)
d-dimensional Hausdorff measure, 425
\(s(\phi)\)
topological entropy of flows, via separated sets, 325
\(s(f)\)
topological entropy, via separated sets, 311
\(w = (w_\alpha)\alpha\)
translation vector, 208
1-parameter group, 68
\(\sigma\)-additive function, 443, 486
\(\sigma\)-algebra, 440
\(Borel, iii, 441\)
generated, 286, 441
up to measure zero, 444
product, 108, 456, 457
\(\sigma\)-finite measure, 8, 75, 442
\(C^0\) topology, 385
\(C^1\) topology, 385
\(C^r\) topology, 470
\(h\)-expansive map, 320, 331, 355
\(k\)-linear form, 473
\(L^2\) convergence, 70
\(L^S\) norm, 480
\(L^p\) norm, 478
\(p\)-integrable function, 478
absolute continuity, 118, 447, 459
theorem, 137
absolutely
continuous measure, 121, 459
summable series, 481
action-angle coordinates, 127
adding machine, 176
additive
function, 442
sequence, 79
adjoint linear operator, 484
admissible sequence, 321
affine function, 292
algebra, 440
compact, 443
generating, 10
of functions, 468
separating, 468
of measures, 241
almost
every point, 99, 455
everywhere, 455
convergence, 455
integrable system, 126
alphabet, 208
alternate form, 473
Anosov
flow, 136
system, 29, 265
approximate eigenvalue, 228
approximation theorem, 444
area form, 207
arithmetic progression, 58
length, 58
atlas
compatible, 470
differentiable, 469
of class \(C^r\), 469
atom, 466, 488
atomic measure, 466
Aubry-Mather set, 133
automorphism
Bernoulli, 282
Kolmogorov, 286, 289
Möbius, 423
of a group, 170
Avogadro constant, 339
Baire space, 122, 124, 471, 475
Banach space, 49, 478, 482
Banach-Alaoglu theorem, 484
Banach-Mazur theorem, 52
barycenter of a measure, 294
basin of a measure, 103, 360
basis
dual, 473
Fourier, 482
Hammel, 483
Hilbert, 482
of neighborhoods, 36, 37, 448
countable, 448
of open sets, 448
countable, 448
of the topology, 448
INDEX

Bernoulli
automorphism, 282
measure, 197, 457
shift, 108, 109, 197
billiard, 138
corner, 138
dispersing, 143
semi-dispersing, 144
table, 138
Birkhoff
ergodic theorem, 66, 71, 73, 75
ergodic theorem for flows, 78
multiple recurrence theorem, 29
normal form theorem, 131, 134
recurrence theorem, 7, 48
Boltzmann
constant, 340
ergodic hypothesis, v, 65, 124
Boltzmann-Sinai ergodic hypothesis, 138
Borel
$\sigma$-algebra, iii, 441
measure, 462
normal theorem, 108
set, 441
Borel-Cantelli lemma, 451
bottom of a pile, 177
boundary of a partition, 265, 349
bounded
distortion, 106, 107, 112
linear functional, 483
linear operator, 484, 485
Bowen-Manning formula, 388, 428
branch (inverse), 362, 429
contracting, 362, 372
Brin-Katok theorem, 270
bundle
cotangent, 129, 473
tangent, 130, 135
Bunimovich
mushroom billiard, 144
stadium billiard, 144
Cantor
set, 425
substitution, 178
Cauchy-Schwarz inequality, 479
Cayley-Klein distance, 423
Chacon
example, 228
substitution, 178
Champernowne constant, 107
change
of coordinates, 469
of variables formula, 304
characteristic function, 449
circle, 16
rotation, 16
class $C^r$
atlas, 469
homeomorphism, 470
manifold, 469
map, 470
closed differential form, 474
coarser
cover, 308
partition, 151, 254
cocycle, 86
cohomological equation, 166
cohomologous potentials, 338, 406
cohomology relation, 338, 343
commuting maps, 29
compact
agebra, 443
group, 173
space, 443
compactness theorem, 41
compatible atlases, 470
complete
measure, 444
measure space, 444
metric space, 465
metrizable space, 471
completely metrizable space, 471
completion of a measure space, 444
complex measure, 445, 486
concave function, 480
condition
Keane, 209
twist, 128, 130, 132, 134
conditional
entropy, 254
expectation, 155, 157, 271
probability, 149
cone, 51, 411
dual, 52, 390
normal, 51
configuration space, 339
conformal
map, 428
repeller, 388, 428
conjecture of entropy, 327
conjugacy (topological), 223, 310
connected space, 469
conservative
flow, 19
map, 19
system, 49, 124
constant
Avogadro, 339
Boltzmann, 340
Champernowne, 107
of expansivity, 267, 319
continued fraction
expansion, 13
INDEX

of bounded type, 120
continuity
  absolute, 118
  at the empty set theorem, 443
  from above theorem, 451
  set of a measure, 37
continuous
  function, 449
  linear functional, 467, 483
  linear operator, 484
  map, 449
contracting inverse branch, 362, 372
contraction, 314
convergence
  almost everywhere, 455
  in $L^2$, 70
  in distribution, 44
  to equilibrium, 410
convex
  function, 480
  hull, 429
  set, 45
convexity, 121, 292
coordinate change, 469
correlation, 187
decay, 215
correlations sequence, 188
cotangent
  bundle, 129, 473
  space, 129, 473
countable basis
  of neighborhoods, 448
  of open sets, 448
countable type shift, 236
countably
  additive function, 442
  generated system, 236
covariance
  matrix, 240
  sequence, 240
cover, 425
corner, 308
diameter, 312, 316, 425
  finer, 308
  open, 308, 443
cross-ratio, 413
cross-section, 91
cube, 446
cylinder, 457
elementary, 457
measurable, 55
open, 55
decay of correlations, 215, 217
decimal expansion, 10
dergodicity, 106
decomposition
  of Oseledets, 87
  theorem of Hahn, 445
  theorem of Lebesgue, 460
degree of a map, 362, 372, 477
Dehn twist, 495
density
  lower, 59
  of a measure, 361, 459
  point, 458
  upper, 59, 61
  zero at infinity, 195
derivation theorem of Lebesgue, 458
derivative, 471, 472
  exterior, 474
  Radon-Nikodym, 361, 459
diagonal, 31
diameter
  of a cover, 312, 316, 425
  of a partition, 266, 461
diffeomorphism, 469, 470
  of class $C^r$, 470
difference (orthogonal), 235
differentiable
  atlas, 469
  manifold, 469
  map, 470
differential
  form, 473, 474
  closed, 474
  exact, 474
dimension
  Hausdorff, 425
  Hilbert, 483
  Diophantine
    number, 168
    vector, 128
Dirac
  mass, 442
  measure, 47, 442
direct sum (orthogonal), 483
discrete
  spectrum, 221, 230
  spectrum theorem, 242
topology, 110
disintegration
  of a measure, 149
  theorem of Rokhlin, 152, 161
dispersing billiard, 143
distance, 462
  associated with Riemannian metric, 476
  Cayley-Klein, 423
  flat on the torus, 105
  hyperbolic, 423
  invariant, 174
  Poincaré, 423
  projective, 412
INDEX

distortion, 106, 107, 112
   lemma, 363
distribution
   function, 44
   Gibbs, 341
divergence of a vector field, 20
domain
   fundamental, 89
   of invertibility, 301
dominated convergence theorem, 456
dual
   basis, 473
   cone, 52, 390
   linear operator, 51, 388
   of a Banach space, 49, 480
   of a Hilbert space, 484
duality, 50, 215, 388, 479
dynamical
   ball, 269, 311
   for a flow, 324
   infinite, 331
   decomposition theorem, 377
   system, iii
eigenvalue, 225
   approximate, 228
   multiplicity of, 225
   elementary cylinder, 457
   elliptic fixed point, 131–133, 135
   generic, 132
   endomorphism
      of a group, 170
      of the torus, 115
      ergodic, 115
   energy
      hypersurface, 125
      of a state, 340
   entropy, v
      conditional, 254
      conjecture, 327
      formula, 405
      formula (Pesin), 280
      function, 265
      Gauss map, 276
      local, 269
      of a communication channel, 251
      of a dynamical system, 257
      of a linear endomorphism, 277
      of a Markov shift, 274
      of a partition, 252
      of a state, 340
      of an open cover, 308
      relative to a partition, 257
      relative to an open cover, 309
      semi-continuity, 265
      topological, 307, 309, 313
   equation
cohomological, 166
   Hamilton-Jacobi, 125, 127, 131
   equidistributed sequence, 180
   equilibrium state, 308, 339, 352
   equivalence
      ergodic, 167, 190, 221, 222, 241
      invariant of, 223
      spectral, 221, 224, 242
      invariant of, 225
      topological, 310
      of flows, 326
   equivalent
      measures, 14, 459
      topologies, 37
   ergodic
      decomposition theorem, 148
      equivalence, 167, 190, 221, 222, 241
      invariant of, 223
      hypothesis, v, 65, 124, 138
      isomorphism, 222, 241
      measure, 75
      system, 97, 98
      theorem
         Birkhoff, 66, 71, 73, 75
         Birkhoff for flows, 78
         Kingman, 66, 80
         Kingman for flows, 87
         multiplicative, 86, 279
         Oseledets, 86, 279
         subadditive, 66, 80
         subadditive for flows, 87
         von Neumann, 66, 69, 75
         von Neumann for flows, 71
         von Neumann multiple, 196
   ergodicity
      decimal expansion, 106
      irrational rotation, 104, 105
      linear endomorphism, 115
      Markov shift, 201
      essential supremum, 480
      essentially bounded function, 480
      Euclidean space, 469
      exact differential form, 474
      exactness, 291
      topological, 366
   example
      Chacon, 228
      Furstenberg, 166
      existence theorem, 35
      for flows, 48
      expanding map, 27
         of the interval, 368
         on a manifold, 360
         on a metric space, 369
      expansive map, 267, 319, 362, 373
         two-sided, 319
      expansivity
strongly affine, 299
uniformly quasi-periodic, 78

functional
bounded, 483
continuous, 467
norm, 467
positive, 454, 467
over a cone, 52
tangent, 53
functions algebra, 468
separating, 468
fundamental domain, 89
Furstenberg
example, 166
theorem, 166
Furstenberg-Kesten theorem, 86

gas
ideal, 140
lattice, 339
Gauss map, 13, 24
entropy, 276
Gaussian
measure, 239
shift, 239, 289
generated
$\sigma$-algebra, 286, 441
topology, 441
generating
algebra, 10
partition, 263, 265
set, 310
for flows, 324
generator
one-sided, 263
two-sided, 263
generated
geodesic, 476
flow, 135, 476
Gibbs
distribution, 341
free energy, 340
state, 340, 342, 357, 387, 389
golden ratio, 9, 186
Grötzsch theorem, 168
Grünewald theorem, 63
Grassmann manifold, 469
Green-Tao theorem, 60
group
1-parameter, 68
automorphism, 170
compact, 173
endomorphism, 170
Lie, 169
linear, 170, 474
locally compact, 170
metricable, 173
orthogonal, 170
special linear, 170, 475
topological, 169
Hölder
function, 217, 409
inequality, 479, 481
map, 463
Haar
measure, 173
theorem, 171
Hahn decomposition theorem, 445, 455
Halmos-von Neumann theorem, 242
Hamilton-Jacobi equation, 125, 127, 131
Hamiltonian
flow, 125, 131
function, 125
non-degenerate, 128
system, 125
vector field, 22, 131
Hammel basis, 483
Hausdorff
dimension, 425
measure, 425
space, 36, 441
hereditary property, 42
heteroclinic
point, 494
Hilbert
basis, 482
dimension, 483
space, 482
Hindman theorem, 168
homeomorphism, 442
twist, 132
homoclinic point, 132
homomorphism of measure algebras, 241
hyperbolic
flow, 135, 476

gas
ideal, 140
lattice, 339
Gauss map, 13, 24
entropy, 276
Gaussian
measure, 239
shift, 239, 289
generated
$\sigma$-algebra, 286, 441
topology, 441
generating
algebra, 10
partition, 263, 265
set, 310
for flows, 324
generator
one-sided, 263
two-sided, 263
generated
geodesic, 476
flow, 135, 476
Gibbs
distribution, 341
free energy, 340
state, 340, 342, 357, 387, 389
golden ratio, 9, 186
Grötzsch theorem, 168
Grünewald theorem, 63
Grassmann manifold, 469
Green-Tao theorem, 60
group
1-parameter, 68
automorphism, 170
compact, 173
endomorphism, 170
Lie, 169
linear, 170, 474
locally compact, 170
metricable, 173
orthogonal, 170
special linear, 170, 475
topological, 169
Hölder
function, 217, 409
inequality, 479, 481
map, 463
Haar
measure, 173
theorem, 171
Hahn decomposition theorem, 445, 455
Halmos-von Neumann theorem, 242
Hamilton-Jacobi equation, 125, 127, 131
Hamiltonian
flow, 125, 131
function, 125
non-degenerate, 128
system, 125
vector field, 22, 131
Hammel basis, 483
Hausdorff
dimension, 425
measure, 425
space, 36, 441
hereditary property, 42
heteroclinic
point, 494
Hilbert
basis, 482
dimension, 483
space, 482
Hindman theorem, 168
homeomorphism, 442
twist, 132
homoclinic point, 132
homomorphism of measure algebras, 241
hyperbolic
flow, 135, 476

gas
ideal, 140
lattice, 339
Gauss map, 13, 24
entropy, 276
Gaussian
measure, 239
shift, 239, 289
generated
$\sigma$-algebra, 286, 441
topology, 441
generating
algebra, 10
partition, 263, 265
set, 310
for flows, 324
generator
one-sided, 263
two-sided, 263
generated
geodesic, 476
flow, 135, 476
Gibbs
distribution, 341
free energy, 340
state, 340, 342, 357, 387, 389
golden ratio, 9, 186
Grötzsch theorem, 168
Grünewald theorem, 63
Grassmann manifold, 469
Green-Tao theorem, 60
group
1-parameter, 68
automorphism, 170
compact, 173
endomorphism, 170
Lie, 169
linear, 170, 474
locally compact, 170
metricable, 173
orthogonal, 170
special linear, 170, 475
topological, 169
Hölder
function, 217, 409
inequality, 479, 481
map, 463
Haar
measure, 173
theorem, 171
Hahn decomposition theorem, 445, 455
Halmos-von Neumann theorem, 242
Hamilton-Jacobi equation, 125, 127, 131
Hamiltonian
flow, 125, 131
function, 125
non-degenerate, 128
system, 125
vector field, 22, 131
Hammel basis, 483
Hausdorff
dimension, 425
measure, 425
space, 36, 441
hereditary property, 42
heteroclinic
point, 494
Hilbert
basis, 482
dimension, 483
space, 482
Hindman theorem, 168
homeomorphism, 442
twist, 132
homoclinic point, 132
homomorphism of measure algebras, 241
hyperbolic
flow, 135, 476

ideal gas, 140
idempotent linear operator, 486
identity

parallelogram, 485
polarization, 485
image of a measure, 45, 50
independent partitions, 252
induced map, 21
inequality

Cauchy-Schwarz, 479
Hölder, 479, 481
Jensen, 480
Margulis-Ruelle, 279
Minkowski, 478, 481
Tchebysheff-Markov, 460
Young, 481
infinite
dynamical ball, 331
matrix, 239
measure, 7
infinitesimal generator, 68
information
function of a partition, 252
of a symbol, 251
of an alphabet, 251
inner product, 479, 482
integer part, 10
integrability, 453
uniform, 88, 460
integrable
function, 453
map, 130
system, 126
integral, 453, 454
first, 125
of a simple function, 453
with respect to a signed measure, 455
with respect to complex measure, 455
interval
d-adic, 498
exchange, 94, 207
irreducible, 209
in \( \mathbb{Z} \), 59
intrinsically ergodic map, 357
invariant
distance, 174
function, 70, 98
measure, 2, 45
of ergodic equivalence, 223
of spectral equivalence, 225
set, 56, 98, 365
inverse branch, 362, 429
contracting, 362, 372
invertibility domain, 301
irrational rotation, 17
ergodicity, 104, 105
irreducible
interval exchange, 209
stochastic matrix, 201
isometrically isomorphic spaces, 467, 483
isometry, 314
linear, 51, 484, 485
isomorphism
ergodic, 222, 241
of measure algebras, 241
iterate of a measure, 45, 50
iterated sum
of a partition, 256, 259
of an open cover, 308, 315, 316
Jacobian, 301
Jacobs theorem, 293, 294
Jensen inequality, 480
Kač theorem, 5
Kakutani-Rokhlin
lemma, 29
tower, 27, 28
Keane
condition, 209
theorem, 210
Kingman ergodic theorem, 66, 80
for flows, 87
Kolmogorov
automorphism, 286, 289
system, 286, 289
Kolmogorov-Arnold-Moser
theorem, 128, 130
theory, v
Kolmogorov-Sinai
entropy, 257
theorem, 261
Koopman operator, 50, 51
lattice
gas, 339
system, 339
state, 339
leaf
stable, 117
unstable, 117
Lebesgue
decomposition theorem, 460
derivation theorem, 458
exterior measure, 446
integral, 453, 454
measurable set, 244, 447, 452
measure, 446
spectrum, 221, 233, 289
rank of, 236, 238
left
invariance, 173
translation, 170
lemma
Borel-Cantelli, 451
distortion, 363
Fatou, 456
Kakutani-Rokhlin, 29
Riemann-Lebesgue, 240
shadowing, 373
Vitali, 459
Zorn, 31
length
of a curve, 476
of an arithmetic progression, 58
INDEX

vector, 208
Levy-Prohorov metric, 39
Lie group, 169
lift
of an invariant measure, 56
of an invariant set, 56
limit of a sequence of sets
inferior, 441
superior, 441
linear
endomorphism of the torus, 115
form, 473
functional
bounded, 483
continuous, 467, 483
norm, 49, 467
positive, 454, 467
positive over a cone, 52
tangent, 53, 357
group, 170, 474
special, 475
isometry, 51, 484, 485
operator
adjoint, 484
bounded, 484, 485
continuous, 484
dual, 51, 388
idempotent, 486
Koopman, 50, 51
normal, 484, 488, 489
positive, 41, 51, 388
positive over a cone, 52
self-adjoint, 484, 486
spectrum, 486
unitary, 484, 488
Liouville
formula, 20
measure, iv, 125
theorem, 20, 21
Lipschitz map, 463
Ljivsic theorem, 387, 406
local
chart, 469
coordinate, 469
diffeomorphism, 477
entropy, 269
local diffeomorphism, 279
locally
compact group, 170
constant function, 215
integrable function, 458
invertible map, 301
logistic map, 319
lower density, 59
Lusin theorem, 464, 466
Lyapunov exponent, 87
Möbius automorphism, 423
manifold
differentiable, 469
Grassmannian, 469
leaf, 137
modeled on a Banach space, 469
of class \( C^r \), 469
Riemannian, 476
stable, 117, 137
symplectic, 129
unstable, 117, 137
Manneville-Pomeau map, 26
map
\( h \)-expansive, 320, 331, 355
conformal, 428
conservative, 19
continuous, 449
decimal expansion, 10
degree, 362, 372, 477
derivative, 471, 472
differentiable, 470
expanding, 27, 369
of the interval, 368
on a manifold, 360
expansive, 267, 319, 362, 373
exponential, 476
first-return, 5, 22, 90, 91
Gauss, 13, 24
Hölder, 463
induced, 24
integrable, 130
interval exchange, 94, 207
intrinsically ergodic, 357
Lipschitz, 463
locally invertible, 301
logistic, 319
Manneville-Pomeau, 26
measurable, 449
minimal, 18, 105
non-degenerate, 130
of class \( C^r \), 470
Poincaré, 90, 91
shift, 60, 109
symplectic, 129
time-1, 5
topologically
exact, 366
mixing, 190
weak mixing, 227
transitive, 110
two-sided expansive, 319
maps
topologically conjugate, 310
topologically equivalent, 310
Margulis-Ruelle
inequality, 279
Markov
measure, 197
shift, 197
entropy, 274
ergodic, 201
finite, 198
mixing, 203
mass distribution principle, 436
Masur-Veech theorem, 211
matrix
hyperbolic, 116
infinite, 239
positive definite, 289
stochastic, 198
aperiodic, 203
irreducible, 201
symmetric, 240
transition, 321
maximal entropy measure, 352
Mazur theorem, 53
mean
information of an alphabet, 251
return time, 6
sojourn time, 65, 71
measurable
cylinder, 55
function, 449
map, 449
partition, 147, 151
set, iii, 440
Lebesgue, 244, 447, 452
space, 440
measure, iv, 442
σ-finite, 8, 75, 442
absolutely continuous, 121
algebra, 241
homomorphism, 241
isomorphism, 241
atomic, 466
barycenter of, 294
basin, 360
Bernoulli, 197, 457
Borel, 462
complete, 444
complex, 445, 486
density, 361
Dirac, 47, 442
ergodic, 75
exterior, 446
finite, 442
flux, 92, 94
Gaussian, 239
Haar, 173
Hausdorff, 425
infinite, 7
invariant, 2, 45
Lebesgue, 446
on the circle, 16
Liouville, iv, 125
Markov, 197
non-atomic, 466
non-singular, 301, 461
of maximal entropy, 352
of probability, iv, 442
physical, 367
positive, 444
product, 108, 456, 457
quotient, 148
reference, 389–391
regular, 462
signed, 50, 445
finite, 445
space, 442
complete, 444
completion, 444
spectral, 486
stationary, 57, 197
suspension of, 93
tight, 465
with finite memory, 196, 206
measures
equivalent, 14, 459
mutually singular, 122, 459
metric
Levy-Prohorov, 39
Riemannian, 475
space, 462
complete, 465
metrizable
group, 173
space, 39, 462
minimal
map, 18, 105
set, 8, 165, 168
system, 163, 165, 210
minimality, 18, 163, 210
minimizing curve, 476
Minkowski inequality, 478, 481
mixing, 188
Markov shift, 203
weak, 226
monkey paradox, 110
monotone
class, 444
theorem, 444
convergence theorem, 455
multiple
natural extension, 57
recurrence theorem
Birkhoff, 29
Poincaré, 29
von Neumann ergodic theorem, 196
multiplicative ergodic theorem, 86, 279
multiplicity
function, 489
INDEX

of a Lyapunov exponent, 87
of an eigenvalue, 225
mushroom billiard, 144
mutually singular measures, 122, 459
natural extension, 54, 57
negative
curvature, 136
part of a function, 453
neighborhood
of a point, 448
of a set, 36
non-atomic measure, 466
non-degenerate
fixed point, 131, 134
Hamiltonian, 128
map, 130
non-lacunary sequence, 29
non-singular measure, 301, 461
non-trivial
partition, 289
probability space, 286
non-wandering
point, 34
super, 63
set, 34
norm, 479, 482
$L^\infty$, 480
$L^p$, 478
of a linear functional, 49, 467
of a matrix, 79
of a measure, 445
of an operator, 174, 323, 326
uniform convergence, 466
normal
cone, 51
linear operator, 484, 488, 489
number, 11, 107, 108
normalized restriction of a measure, 147, 163
number
Diophantine, 168
normal, 11, 107, 108
odometer, 176
one-sided
expansivity, 267
generator, 263
iterated sum
of a partition, 259
of an open cover, 308, 316
shift, 109
open
cover, 308, 443
diameter, 312, 316
cylinder, 55
operator
dual, 51, 388
Koopman, 50, 51
norm, 174, 323, 326
normal, 484, 488, 489
positive, 41, 51, 388
over a cone, 52
Ruelle-Perron-Frobenius, 388
transfer, 215, 388
orbital
average, 73
sum, 332, 388
orthogonal
complement, 67, 483
difference, 235
direct sum, 483
group, 170
projection, 66
vectors, 482
orthonormal set, 482
Oseledepts
decomposition, 87
ergodic theorem, 86, 279
filtration, 86
Oxtoby-Ulam theorem, 124
parallelogram identity, 485
part
fractional, 10
integer, 10
negative, 453
positive, 80, 453
partition, 6, 251, 461
boundary, 265, 349
defined by a cover, 40, 266
diameter, 266, 461
generating, 263, 265
measurable, 147, 151
non-trivial, 289
of $Z$, 58
partitions
coarser, 151, 254
finer, 151, 254
independent, 252
path connected space, 477
periodic pre-orbit, 374
permutation, 77
Perron-Frobenius theorem, 198
Pesin entropy formula, 280, 405
phase transition, 339
physical measure, 367
pile
bottom, 177
simple, 177
top, 177
piling method, 177
Poincaré
distance, 423
INDEX

first-return map, 90, 91
last theorem, 132
recurrence theorem, 4, 7
for flows, 5
multiple, 29
Poincaré-Birkhoff fixed point theorem, 132
point
heteroclinic, 494
non-wandering, 34
of density, 458
recurrent, 7
simultaneously recurrent, 29
super non-wandering, 63
transverse homoclinic, 132
pointwise topology, 44
polarization identity, 485
Portmanteau theorem, 37
positive
definite matrix, 239
linear functional, 454, 467
linear operator, 41, 51, 388
measure, 444
over a cone, 52
part of a function, 80, 453
potential, 307, 332
cohomologous, 338, 406
pre-orbit, 55, 374
periodic, 374
pressure, 307, 332
of a state, 340, 341
primitive substitution, 178
principle
least action (Maupertuis), 340
mass distribution, 436
variational, 340, 344
probability
conditional, 149
measure, iv, 442
space, 442
non-trivial, 286
standard, 241
transition, 196
product
σ-algebra, 108, 456, 457
inner, 479, 482
measure, 108, 457
of measures, 456
countable case, 457
finite case, 456
space, 456, 457
topology, 110, 451, 458
Prohorov theorem, 42
projection, 486
orthogonal, 66
stereographic, 469
projective
distance, 412
quotient, 412
space, 477
pseudo-orbit, 373
periodic, 373
 quasi-periodic function, 127
quotient
measure, 148
projective, 412
Radon-Nikodym
derivative, 361, 459
theorem, 459
random variable, 44
rank of Lebesgue spectrum, 236, 238
rational rotation, 17
rationally independent vector, 18, 209
Rauzy-Veech renormalization, 214
rectangle, 118, 446
recurrent point, 7
simultaneously, 29
reference measure, 389–391
regular
measure, 462
value of a map, 474
renormalization of Rauzy-Veech, 214
repeller, 427
conformal, 388, 428
residual set, 124, 471, 475
return
first, 5
simultaneous, 30
time, 89, 90
time (mean), 6
Riemann sum, 454
Riemann-Lebesgue lemma, 240
Riemannian
manifold, 476
metric, 475
submanifold, 476
Riesz-Markov theorem, 445, 467
right
invariance, 173
translation, 170
Rokhlin
disintegration theorem, 152, 161
formula, 301, 303, 368
root of a system, 291
rotation, 16
irrational, 17
number, 132
on the circle, 16
on the torus, 18
rational, 17
spectrum, 232
Ruelle
inequality, 279
Ruelle theorem, 342, 387
Ruelle-Perron-Frobenius operator, 388

Sard theorem, 475
Schauder-Tychonoff theorem, 45
semi-continuity of the entropy, 265
semi-continuous function, 33
self-adjoint linear operator, 484, 486
semi-dispersing billiard, 144
separable Hilbert space, 483
space, 39, 464, 467
separated set, 311
for flows, 324
separating functions algebra, 468
sequence, 243
sequence additive, 79
admissible, 321
equidistributed, 180
non-lacunary, 29
of correlations, 188
of covariance, 240
separating, 243
subadditive, 79, 80
series absolutely summable, 481
Fourier, 105, 115
set Aubry-Mather, 133
Borel, 441
Cantor, 425
covex, 45
generating, 310
for flows, 324
invariant, 56, 98, 365
Lebesgue measurable, 244, 447, 452
measurable, 440
minimal, 8, 165, 168
non-wandering, 34
of continuity of a measure, 37
of invariant vectors, 67
orthonormal, 482
residual, 124, 471, 475
separated, 311
for flows, 324
strongly convex, 294
syndetic, 9, 168
tight, 42
transitive, 122
with zero volume, 475
shadowing lemma, 373
Shannon-McMillan-Breiman theorem, 269
shift Bernoulli, 108, 109, 197
Gaussian, 239, 289
map, 60, 109
Markov, 197
entropy, 274
ergodic, 201
finite, 198
mixing, 203
multi-dimensional, 339
of countable type, 236
of finite type, 199, 321
one-sided, 109
two-sided, 60, 109
Sierpinski triangle, 437
signed measure, 445
finite, 445
simple function, 450
pale, 177
simultaneous return, 30
simultaneously recurrent point, 29
Sinai ergodicity theorem, 143
Sinai, Ruelle, Bowen theory, v
skew-product, 54
space 
  Baire, 122, 124, 471, 475
  Banach, 49, 478, 482
  compact, 443
  completely metrizable, 471
  connected, 469
  cotangent, 129, 473
dual, 49, 480, 484
  Euclidean, 469
  Hausdorff, 36, 441
  Hilbert, 482
  separable, 483
  Lebesgue, 241, 244, 246
  measurable, 440
  measure, 442
  complete, 444
  metric, 462
  complete, 465
  metrizable, 39, 462
  complete, 471
  of configurations, 339
  path connected, 477
  probability, 442
  non-trivial, 286
  product, 456, 457
  projective, 477
  separable, 39, 464, 467
tangent at a point, 130, 471
topological, 441
topological vector, 45
spaces
  isometric, 467, 483
  isomorphic, 467, 483
  special linear group, 170, 475
specification, 381, 382
by periodic orbits, 382
spectral
equivalence, 221, 224, 242
invariant of, 225
gap property, 215, 422
measure, 486
radius, 52, 323
representation theorem, 489
theorem, 488
spectrum
discrete, 221, 230
Lebesgue, 221, 233
rank of, 236, 238
of a linear operator, 225, 486
of a rotation, 232
of a transformation, 225
sphere of dimension \(d\), 469
spin system, 339
stable
foliation, 116, 117, 137
leaf, 117, 137
manifold, 117, 137
set, 55
stadium billiard, 144
standard probability space, 241
state
energy of, 340
entropy of, 340
equilibrium, 308, 329, 352
Gibbs, 340, 342, 357, 387, 389
of a lattice system, 339
pressure of, 340, 341
stationary measure, 57, 197
stereographic projection, 469
stochastic matrix, 198
aperiodic, 203
irreducible, 201
Stone theorem, 68
Stone-Weierstrass theorem, 468
stronger topology, 37
strongly
affine function, 299
convex set, 294
subadditive
ergodic theorem, 66, 80
for flows, 87
sequence, 79, 80
subcover, 308, 443
submanifold, 470
Riemannian, 476
substitution, 177, 179
Cantor, 178
Chacon, 178
Feigenbaum, 178
Fibonacci, 178, 319
primitive, 178
Thue-Morse, 178
sum
direct orthogonal, 483
of a family of subspaces, 482
of a family of vectors, 482
of open covers, 308
of partitions, 252
orbital, 332, 388
Riemann, 454
super non-wandering point, 63
support
of a measure, 448
of a spectral measure, 486
suspension
flow, 89
of a measure, 90, 93
of a transformation, 89
symmetric
difference, 444
matrix, 240
symplectic
form, 129
manifold, 129
map, 129
syndetic set, 9, 168
system
almost everywhere invertible, 248
almost integrable, 126
Anosov, v
aperiodic, 29, 265
conservative, 49, 124
countably generated, 236
ergodic, 97, 98
Hamiltonian, 125
integrable, 126
Kolmogorov, 286, 289
lattice, 339
state, 339
Lebesgue spectrum
rank of, 236, 238
minimal, 163, 165, 210
mixing, 188
root, 291
spin, 339
totally dissipative, 49
uniquely ergodic, 163
weak mixing, 191, 226
with discrete spectrum, 221, 230
with finite memory, 196, 206
with Lebesgue spectrum, 221, 233, 289
Szemerédi theorem, 60, 61
tangent
bundle, 130, 135, 472
unit, 135, 477
linear functional, 53, 357
space at a point, 130, 471
INDEX

Tchebysheff-Markov inequality, 460

theorem
absolute continuity, 137
Anosov, 136
approximation, 444
Banach-Alaoglu, 50, 484
Banach-Mazur, 52
Birkhoff
ergodic, 66, 71, 73, 75
ergodic for flows, 78
multiple recurrence, 29
normal form, 131, 134
recurrence, 7, 48
Borel normal, 108
Brin-Katok, 270
compactness, 41
continuity
at the empty set, 443
from above, 451
from below, 451
discrete spectrum, 242
disintegration, 152
dominated convergence, 456
dynamical decomposition, 377
ergodic decomposition, 148
existence of invariant measures, 35
for flows, 48
extension of measures, 443
Friedman-Ornstein, 290
Furstenberg, 166
Furstenberg-Kesten, 86
Gottschalk, 168
Grünewald, 63
Green-Tao, 60
Haar, 171
Hahn decomposition, 445
Halmois-von Neumann, 242
Hindman, 168
Jacobs, 293, 294
Kač, 5
Keane, 210
Kingman ergodic, 66, 80
for flows, 87
Kolmogorov-Arnold-Moser, 128, 130
Kolmogorov-Sinai, 261
Lebesgue
decomposition, 460
derivation, 458
Liouville, 20, 21
Livšic, 387, 496
Lusin, 464, 466
Masur-Veech, 211
Mazur, 53
monotone class, 444
monotone convergence, 455
multiplicative ergodic, 86, 279
Oseledec, 86, 279
Oxtoby-Ulam, 124
Perron-Frobenius, 198
Poincaré
multiple recurrence, 29
recurrence, 4, 7
Poincaré-Birkhoff fixed point, 132
Portmanteau, 37
Prohorov, 42
Radon-Nikodym, 459
Riesz-Markov, 445, 467
Rokhlin, 152, 161
Ruelle, 342, 387
Sard, 475
Schauder-Tychonoff, 45
Shannon-McMillan-Breiman, 260
Sinai ergodicity, 143
spectral, 488
spectral representation, 489
Stone, 68
Stone-Weierstrass, 468
subadditive ergodic, 66, 80
for flows, 87
Szemerédi, 60, 61
Tychonoff, 110
van der Waerden, 58, 60
von Neumann ergodic, 66, 69, 75
for flows, 71
multiple, 196
Weyl, 180
Whitney, 476
Thue-Morse substitution, 178
tight
measure, 465
set of measures, 42
time
average, 73
constant of a subadditive sequence, 88
mean sojourn, 65, 71
of first return, 5, 23, 91
of return, 89, 90
time-1 map, 5
top of a pile, 177
topological
conjugacy, 223, 310
entropy, 307, 309, 313
for flows, 325, 326
equivalence, 310
of flows, 326
factor, 310, 386
group, 169
space, 441
vector space, 45
weak mixing, 227
topologically
conjugate maps, 310
equivalent maps, 310
exact map, 366