Texts and Readings in Mathematics 79

Anima Nagar
Riddhi Shah
Shrihari Sridharan *Editors*

Elements of Dynamical Systems

HINDUSTAN BOOK AGENCY

Springer

Texts and Readings in Mathematics

Volume 79

The *Texts and Readings in Mathematics* series publishes high-quality textbooks, research-level monographs, lecture notes and contributed volumes. Undergraduate and graduate students of mathematics, research scholars and teachers would find this book series useful. The volumes are carefully written as teaching aids and highlight characteristic features of the theory. Books in this series are co-published with Hindustan Book Agency, New Delhi, India.

Anima Nagar · Riddhi Shah · Shrihari Sridharan
Editors

Elements of Dynamical Systems

Lecture Notes from NCM School

HINDUSTAN BOOK AGENCY

Springer

Editors
Anima Nagar
Department of Mathematics
Indian Institute of Technology Delhi
New Delhi, India

Riddhi Shah
School of Physical Sciences
Jawaharlal Nehru University
New Delhi, India

Shrihari Sridharan
School of Mathematics
Indian Institute of Science Education
and Research - Thiruvananthapuram
Thiruvananthapuram, India

ISSN 2366-8717 ISSN 2366-8725 (electronic)
Texts and Readings in Mathematics
ISBN 978-981-16-7962-9 (eBook)
https://doi.org/10.1007/978-981-16-7962-9

This work is a co-publication with Hindustan Book Agency, New Delhi, licensed for sale in all countries in electronic form only. Sold and distributed in print across the world by Hindustan Book Agency, P-19 Green Park Extension, New Delhi 110016, India.
Jointly published with Hindustan Book Agency ISBN of the Hindustan Book Agency edition: 978-93-86279-83-5

Mathematics Subject Classification: 37-XX, 11-XX, 37Axx, 37Bxx, 37B10

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Various Advanced Training in Mathematics (ATM) schools, originally launched by the National Board for Higher Mathematics (NBHM), are the most successful workshops which have helped students, teachers and researchers to enhance their scholarship and improve research. The efforts of the National Centre of Mathematics (NCM) and its apex committee in continuing this yeoman service by conducting several Annual Foundation Schools (AFS), Advanced Instructional Schools (AIS), NCM Workshops (NCMW), Instructional Schools for Teachers (IST) and Teachers' Enrichment Workshops (TEW) throughout the year are praiseworthy. Organised by the NCM, with support from NBHM, Department of Atomic Energy (DAE), Government of India, these workshops have attained a status of their own amidst the Math community, within the country and globally.

These lecture notes grew out of a three-week AIS on "Ergodic Theory and Dynamical Systems" (https://www.atmschools.org/2017/ais/etds) that was conducted at the Indian Institute of Technology Delhi (IITD), during 4th–23rd December 2017, organised by NCM, with the support of NBHM, DAE, Government of India. The speakers at this school were C. S. Aravinda, Siddhartha Bhattacharya, S. G. Dani, Anish Ghosh, V. Kannan, Anima Nagar, C. R. E. Raja and Kaushal Verma. Their lectures were aided by the huge support of the tutors of the programme, Nikita Agarwal, P. Chiranjeevi, Manoj Choudhuri, Rajkumar Krishnan, Shrihari Sridharan and Puneet Sharma. We are thankful to the contributions of all the lecturers and the tutors.

Dynamics is the study of the evolution of any given system with time, governed by some physical law. Different laws imposed on the system could give rise to a variety of dynamical systems. The laws may arise in a variety of ways; some with respect to the structure of the underlying space where the system is manifested, some with respect to nature of the action on the space, some with respect to our notion of observation of the evolution etc. The topic of dynamical systems is thus very rich; with different researchers focussing on different aspects.

These lecture notes are intended to help a new researcher understand various aspects of dynamical systems. In keeping with the true spirits of the availability of a variety of means to study dynamical systems, this book begins with chapters on various kinds of dynamics; real dynamics, topological dynamics, ergodic theory,

symbolic dynamics, complex dynamics. Further, as is natural for a topic that spans a variety of interests across areas, the theory of dynamical systems has useful applications across a broad spectrum of areas in mathematics, such as topology, complex analysis, number theory and representation theory. In this book, we later provide a glimpse of such applications to number theory and game theory.

Every chapter of this book specialises in one aspect of dynamical systems; and thus begins at an elementary level and goes on to cover fairly advanced materials. Even though the lectures were delivered to a slightly mature audience comprising of graduate students from across the country, the chapters have been written by the respective authors so articulately that a beginner can read and understand the materials covered, with a bit of an effort.

In the first chapter, we study dynamics of maps on the real line or on an interval there. Most of the theorems proved in this chapter are special to the real line; their analogues do not hold in general dynamical systems. Section one starts with the definitions of such terms as fixed point, periodic point, eventually periodic point, recurrent point and non-wandering point. Their inter-relations are observed. Next, elementary examples of maps like contraction map, identity map, squaring map, tent map, logistic map and shift map are introduced. In each of these examples, periodic points, recurrent points, etc. are explicitly calculated. Which kinds of subsets can arise as the set $\mathrm{Fix}(f)$ of all fixed points? This and four of its analogues are answered. Three more notions, namely invariant sets, omega-limit sets and cycles are introduced.

In section two, notions of attracting and repelling cycles are studied. It starts with the classical theorem of Banach known as the *contraction mapping theorem.* There are various ways of understanding the attracting nature of a fixed point, from the view points of calculus, topology, metric, etc. We discuss mutual implications among them. Several counter examples are provided to disprove some of the implications.

In section three, topological transitivity is studied through various equivalent formulations. Five different proofs are included for the fact that the tent map is topologically transitive. These proofs lead to five different general theorems that open up five significant directions of study. Incidentally, some more concepts such as topological conjugacy, Markov maps and expanding maps are also introduced. In section four, Devaney's definition of chaos is introduced through three ingredient properties. The independence of these three is established by a set of eight counter examples. While doing so, about a dozen propositions involving transitivity, sensitivity and dense periodicity are proved. In section five, it is seen that the independence results obtained in the previous section are not valid when the underlying space is restricted. For example, we prove that on the real line, every transitive map is necessarily chaotic.

Section six is mainly devoted to a theorem of Sarkovskii on cycle lengths available for real maps and the forcing relation among them. Here the proofs are merely outlined. Next comes a short section in which Baire Category theorem and another theorem (that have been used earlier) are proved. The chapter ends with a short section consisting of notes and exercises.

In chapter two, we introduce G-systems and describe basic notions such as recurrence, minimality and enveloping semigroups. We provide a proof of Van der Waerden's theorem. We also discuss proximal and distal notions and its relation with

enveloping semigroup. Topological dynamics is inspired by the qualitative study of differential equations, initiated by the approach of Henri Poincare, and followed largely by the contribution of G. D. Birkhoff.

G-systems are jointly continuous actions of a topological group on a Hausdorff space. This abstract approach was initiated by W. H. Gottschalk and G. A. Hedlund. We adopt their approach to study the basic notions of recurrence in the first section. The second section is devoted to minimal systems which is fundamental to many recurrence theorems. Such phenomena have a wide range of applications and we provide one such application in the field of number theory. We discuss the famous proof of the celebrated Van der Waerden's theorem given by H. Furstenberg and B. Weiss in the third section. In the fourth section, we discuss the algebraic theory of enveloping semigroups that form a fundamental tool to study topological dynamics. The notions of proximal and distal systems are important aspects, which we discuss in the fifth section. The sixth section is basically dedicated to the evergreen notion of topological transitivity and its various forms.

Chapter three is a gently paced introduction to some of the key ideas in the general topic of Ergodic theory, providing essential background to discuss some of the cornerstone results in the field.

The first couple of decades of the twentieth century witnessed a definitive, neat and clear understanding of the all important notions of measure in a general context. Apart from serving as a warm up on the rudiments of measure, the second section of this chapter intends to particularly highlight the work of C. Carathéodary in this context, and point out the possible logic behind the introduction of the Carathéodory criterion for a set to be measurable. The section ends with a quick description of Hausdorff measures and Hausdorff dimension.

Recalling a motivation from certain questions in statistical mechanics, the main aim of the third section is to give a proof of the celebrated Birkhoff ergodic theorem. Also known as the pointwise ergodic theorem, first proved in 1931 by G. D. Birkhoff, this lofty result brought in much clarity on the notion of ergodicity, and triggered significant progress in the mathematical aspects of the theory.

Building further on the discussion in the previous sections, the fourth and final section sketches the proof of ergodicity of one of the earliest interesting examples of an ergodic dynamical system—the geodesic flow on the unit tangent bundle of closed surface of constant negative curvature. First proven in the year 1934 by G. Hedlund, the proof sketched here is the one due to E. Hopf which has inspired monumental later work in hyperbolic dynamics. The first subsection to section four may also serve as an introduction to hyperbolic geometry. Thus, the third chapter essentially captures the spirit of the remarkable development heralding the beginnings of this important area of research during the first four decades of the twentieth century.

Symbolic dynamics is the study of shift spaces, which consist of infinite or bi-infinite sequences on a pre-determined alphabet set. These sequences almost capture the essence of abstract systems and provide a simplified model of study. Codings give mappings between two such shift spaces. Further, aided by the combinatorial, algebraic, topological and measure-theoretic invariants, codings give a subtle description of many dynamical properties, as well.

After introducing the setup of symbolic dynamics in the first section of chapter four, we discuss some basic properties in the second section. The concept of entropy is defined in the third section, and the fourth section deals with methods to compute such entropy. In the fifth section, a class of symbolic dynamical systems related to tiling spaces is defined and a profound result due to M. Szegedy is proved. The last section is devoted to an algebraic dynamical system known as 3-dot system, which is used to study symbolic systems that can exhibit strong rigidity property.

The purpose of chapter five is to present some basic ideas and tools in complex dynamics. Starting with some elementary observations that motivate us to study this topic in detail, we make use of the various versions of Montel's theorem that describes normality in a family of holomorphic functions defined on a domain in the Riemann sphere, $\mathbb{P}^1 = \mathbb{C} \cup \{\infty\}$. Dichotomising the Riemann sphere using Montel's normality criterion on the family of iterates of a rational map, we obtain the Fatou and Julia sets of the considered rational map. Various properties of these two sets are then investigated; one important property being the non-vacuousness of the Julia set for rational maps of degree at least 2. Answering our natural curiosity about a similar property for the Fatou set of a rational map, we construct a family of rational maps for which the Julia set is all of $\mathbb{P}^1$; which implies that the Fatou set is empty, in this case. Lattès' example is a simple case of this construction.

The authors then focus on some statements that characterise the Julia set of a rational map, alternatively using various results from complex analysis. These statements are more useful in determining the Julia sets computationally. Then, the focus shifts to studying local normal forms near fixed points and the classification of Fatou components for rational maps.

One important result in this field pertains to the relation between the dense set of pre-images of any generic point in the Julia set and the equilibrium measure of any compact subset of $\mathbb{P}^1$, using the energy integral, as encapsulated by a result due to H. Brolin. The authors build their case for the Brolin's theorem in $\mathbb{P}^1$ and discuss analogous results in higher dimensions.

Recent decades have seen dramatic progress in the study of ergodic aspects of group actions on homogeneous spaces of Lie groups. Much of this progress, beginning with Margulis' famous proof of Oppenheim's conjecture, has been closely associated to Diophantine analysis. Another, more recent example is the important work of Einsiedler, Katok and Lindenstrauss towards Littlewood's conjecture. The aim of chapter six is to present some topics at the interface of homogeneous dynamics and number theory with the aim of giving the reader a glimpse of the rich connections between the two subjects. The goal is to whet the appetite of the reader. The interested reader can then move on to a more systematic and detailed source like the book by Einsiedler and Ward. This is suitable for talented undergraduates with some background in Lie groups, for graduate students, as well as for mathematicians who wish to get acquainted with the area.

The aim of chapter seven is to give an introduction to a notion of "large subsets" of Euclidean and other similar spaces, that has attracted much attention in the recent decades, in the theory of Diophantine approximation, geometry, and dynamics of flows on homogeneous spaces. The sets are defined in terms of existence of winning

strategies for various two-player infinite games. Their origin goes back to a 1966 paper of W. M. Schmidt, which made interesting observations about the set of badly approximable real numbers having certain unusual largeness properties, which has now found generalisations in a variety of contexts.

As a specialist may have observed, the topics dealt with in each of these chapters is an area of research in its own right, however that depends on the other areas also described in the other chapters. One may religiously cover all materials in this book, if one is interested to give a year-long course on various elements of dynamical systems, as the title of the book suggests. However, within the book lies various ideas for a one-semester course; each of the combination below describing one such.

- chapters 1–3 and 5;
- chapters 1, 2, 4 and 5;
- chapters 2–4 and 5;
- chapters 3, 6 and 7; etc.

We are grateful to the participants Mahboob Alam, G. K. Chaitanya, Haritha Cheriyath, Pramod Das, Shreyasi Datta, Mukta Garg, Dileep Kumar, Dinesh Kumar, Pabitra Narayan Mandal, Manoj B. Prajapati, Yogesh Prajapaty, Manish Rajput, Manpreet Singh, Pradeep Singh, Sharvari Neetin Tikekar and Atma Ram Tiwari, for their special efforts in taking notes which helped the authors in preparing their lecture notes. It is a pleasure to thank the students for their contributions to these lecture notes.

Lastly, we thank the Indian Institute of Technology Delhi (IITD) for their excellent hospitality.

New Delhi, India — Anima Nagar
New Delhi, India — Riddhi Shah
Thiruvananthapuram, India — Shrihari Sridharan

Contents

Real Dynamics

V. Kannan

1 Introduction and Preliminaries from Topological Dynamics

1.1 Introduction

Real dynamics is the study of those discrete dynamical systems for which the underlying set (called the phase space) is the real line $\mathbb{R}$ or the unit interval $I = [0, 1]$, or occasionally some other subset of $\mathbb{R}$. But the definitions will be given in a more general setting. Most of the examples will be given from real dynamics. Other examples are also provided to see the contrast with real dynamics.

1.2 Preliminaries

Let $\mathbb{N}$ denote the set of all positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let $\mathbb{Q}$ denote the set of all rational numbers.

A dynamical system is a pair (X, f) where X is a topological space and f is a continuous map from X to X. The composition $f \circ f$ will be denoted by f^2. Recursively f^n denotes the n-fold composition of f, for every positive integer n. By convention, f^0 is the identity map.

The sequence $(f^n(x))_{n=0}^{\infty}$ is called the f-trajectory of x in X. Its set

$$\left\{y \in X \mid y = f^n(x) \text{ for some nonnegative integer } n\right\}$$

V. Kannan (✉)
SRM University, Amaravati, AP, India
e-mail: kannan.v@srmap.edu.in; vksm.uoh@nic.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_1

is called the orbit of x and it is denoted by $O(x, f)$. Dynamics is the study of eventual behaviour of the trajectories in a dynamical system.

Here are some dynamical properties of points and these will be defined below.

- Fixed point
- Periodic point
- Eventually fixed point
- Eventually periodic point
- Recurrent point
- Non-wandering point

Definition 1.1 1. An element $x \in X$ is said be a fixed point of (X, f) if $f(x) = x$.
2. An element $x \in X$ is said to be a periodic point of (X, f) if $f^n(x) = x$ for some positive integer n.
3. If $x \in X$ is a periodic point of (X, f), its period is the smallest positive integer n such that $f^n(x) = x$. (In particular, a fixed point is a periodic point of period 1.)
4. An element $x \in X$ is said to be an eventually fixed point if $\exists n \in \mathbb{N}$ such that $f^n(x)$ is a fixed point.
5. An element $x \in X$ is said to be an eventually periodic point if $f^n(x)$ is a periodic point for some $n \in \mathbb{N}$.
6. An element $x \in X$ is called a recurrent point of (X, f) if for every neighbourhood V of x, $\exists n \in \mathbb{N}$ such that $f^n(x) \in V$.
7. An element $x \in X$ is a non-wandering point of (X, f) if for every neighbourhood V of x, $\exists y \in V$ and $\exists n \in \mathbb{N}$ such that $f^n(y)$ is also in V.

We use the following notations:

$$\begin{aligned} Fix(f) &= \text{Set of all fixed points of } f. \\ P(f) &= \text{Set of all periodic points of } f. \\ EP(f) &= \text{Set of all eventually periodic points of } f. \\ R(f) &= \text{Set of all recurrent points of } f. \\ \Omega(f) &= \text{Set of all non-wandering points of } f. \end{aligned}$$

Proposition 1.2 $Fix(f) \subset P(f) \subset R(f) \subset \Omega(f)$.

Proof 1. Fixed points are precisely the periodic points of period 1.
2. If p is a periodic point of period n, and if V is a neighbourhood of p, then $f^n(p) = p \in V$ and thus p is a recurrent point.
3. Let p be a recurrent point and let V be a neighbourhood of p. Then $\exists n \in \mathbb{N}$ such that $f^n(p) \in V$. Now p and $f^n(p)$ in V are as required in the definition of a non-wandering point. So p is non-wandering. □

Example 1.3 *(Contraction map)* Let $f(x) = \dfrac{x}{2}$ on I. Then it is clear that $f^n(x) = \dfrac{x}{2^n}$ $\forall n \in \mathbb{N}$. Every trajectory is decreasing to 0 and 0 is the only fixed point. Moreover, we have in this case that

$$Fix(f) = P(f) = R(f) = \Omega(f) = \{0\}.$$

Illustration: The equation $\frac{x}{2} = x$ has only one solution, namely $x = 0$. Therefore $Fix(f) = \{0\}$. Let $n \in \mathbb{N}$. The equation $\frac{x}{2^n} = x$ has only one solution, namely $x = 0$. Therefore, $P(f) = \{0\}$.

If $x > 0$, then the open interval $\left(\frac{3x}{4}, \frac{3x}{2}\right) \cap I$ is a neighbourhood of x that contains no other element of the orbit of x. Therefore x is not recurrent. If y is any element in this open interval, then $\frac{y}{2}$ is below and outside it, and we can prove that no other element from the orbit of y belongs to this interval. So, $R(f) = \{0\} = \Omega(f)$. □

Example 1.4 *(Identity map)* Let X be any topological space. Let f be the identity map on X. Then

$$Fix(f) = P(f) = R(f) = \Omega(f) = EP(f) = X.$$

Example 1.5 *(The squaring map on* $\mathbb{R}$*)* Let $f(x) = x^2 \ \forall x \in \mathbb{R}$.

Illustration: $Fix(f) = \{0, 1\}$ because the equation $x^2 = x$ has two solutions, namely $x = 0$ and $x = 1$. All trajectories are eventually monotonic; some are strictly increasing; some are strictly decreasing; some are constant. A closer look gives that there are five kinds of trajectories namely:

1. Constant sequence like $\{1, 1, 1, \ldots\}$.
2. Strictly increasing sequence like $\{2, 4, 16, \ldots, \}$ diverging to ∞.
3. Strictly decreasing sequence like $\left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \ldots\right\}$ converging to 0.
4. Eventually constant sequence like $\{-1, 1, 1, \ldots\}$ converging to 1.
5. Non-monotonic but eventually monotonic sequence like $\left\{-\frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \ldots\right\}$ converging to 0.

From this we can prove that $EP(f) = \{-1, 0, 1\}$ and

$$P(f) = Fix(f) = R(f) = \Omega(f).$$ □

Example 1.6 *(The Tent Map)* Let $f : I \longrightarrow I$ be defined by

$$f(x) = \begin{cases} 2x & \text{if } x \leq \frac{1}{2} \\ 2 - 2x & \text{if } x > \frac{1}{2}. \end{cases}$$

Illustration: This is called a tent map because its graph looks like a tent with the point $\left(\frac{1}{2}, 1\right)$ in the roof and with the points $(0, 0)$ and $(1, 0)$ on the ground.

$Fix(f) = \{0, \frac{2}{3}\}$ is obtained by solving $2x = x$ and $2 - 2x = x$ separately. One can also verify that the trajectory of $\frac{2}{5}$ is $\frac{2}{5}, \frac{4}{5}, \frac{2}{5}, \frac{4}{5}, \ldots$. Thus, both $\frac{2}{5}$ and $\frac{4}{5}$ are periodic points of period 2.

By elementary but clever methods, the following have been proved and are available in some books (for example, in [8]).

$$\begin{aligned} P(f) &= \left\{ \frac{2m}{2n+1} \in I \mid 0 \leq m \leq n \in \mathbb{N}_0 \right\}. \\ EP(f) &= \{\text{All rational numbers in } I\}. \\ R(f) &= \text{An uncountable dense set with a dense complement.} \\ \Omega(f) &= \text{The whole set } I. \end{aligned}$$

Thus in this example, these five sets are distinct. Some parts of these results will be proved in a later section. □

Example 1.7 *(The Logistic maps* [10]*)* For each $\mu > 0$, the map $x \longmapsto \mu x(1-x)$ from $\mathbb{R}$ to $\mathbb{R}$ is called a logistic map. When $\mu = 4$, it takes I onto I. For different values of μ, these maps may have different dynamical properties. Its fixed points are 0 and $1 - \frac{1}{\mu}$.

Example 1.8 *(The Shift map)* Let Σ_2 be the set of all (one-sided) sequences of 0's and 1's. For each word w over the alphabet set $\{0, 1\}$, let $V_w = \{x \in \Sigma_2 \mid w \text{ is a prefix of } x\}$.

Illustration: Note that w has finite length k, whereas x has infinitely many terms. We say that w is a prefix of x if $w_i = x_i$ holds for all $i \leq k =$ length of w. Here w_i denotes the i^{th} symbol in the word w; similarly x_i denotes the i^{th} term in the sequence x.

We now use these sets V_w to define a topology on the uncountable set Σ_2. It is that topology for which the family $\{V_w \mid w \text{ is a word over } \{0, 1\}\}$ is a base. Equivalently, it is the same as the product topology, when Σ_2 is regarded as the product $\{0, 1\} \times \{0, 1\} \times \cdots$. It can also be described in terms of a metric on Σ_2, but we now omit this description.

Next, we define $\sigma : \Sigma_2 \longrightarrow \Sigma_2$ by the rule $(\sigma(x))_n = x_{n+1}$, $\forall n \in \mathbb{N}$ and $\forall x \in \Sigma_2$. This means that σ shifts the sequence x by one position to its left side. We can prove that σ is a continuous map from Σ_2 to itself.

$Fix(\sigma)$ is a set that has two elements $\overline{0}$ and $\overline{1}$. Here $\overline{0}$ denotes the constant sequence $\{0\,0\,0\cdots\}$ and $\overline{1}$ has a similar meaning.

If w is any word over $\{0, 1\}$, then $\overline{w}$ denotes the sequence $\{w\,w\,w\,\cdots\}$ obtained by concatenating infinite number of w's. If k is the length of w, it is easy to see that $\sigma^k(w) = w$. Thus $\overline{w}$ is a periodic point for the shift map. One can also prove that there are no other periodic points. In this manner, there are exactly 2^n periodic points x satisfying $\sigma^n(x) = x$.

$$P(\sigma) = \{\overline{w} \mid w \in \{0, 1\}\}$$

is a countably infinite dense subset of Σ_2. It is dense because, if V_w is any basic open set, there is an element of $P(\sigma)$ there, namely $\overline{w}$. If u and v are two words over $\{0, 1\}$, then $u\overline{v}$ (obtained by prefixing u to the sequence $\overline{v}$) is an eventually periodic point. $EP(\sigma)$ consists precisely of such points, and is therefore another countable dense set. $R(\sigma)$ contains $P(\sigma)$ strictly; in fact it is uncountable. $\Omega(\sigma)$ is the whole Σ_2. □

Proposition 1.9 *$x \in EP(f)$ if and only if the orbit of x is finite.*

Proof Let x be eventually periodic. Then $\exists k \in \mathbb{N}$ such that $y = f^k(x)$ is periodic. And $\exists n \in \mathbb{N}$ such that $f^n(y) = y$. Now every $f^m(x)$ is of the form $f^i(x)$ for some $i < k + n$. [In particular $f^{k+n}(x) = f^k(x)$; $f^{k+n+1}(x) = f^{k+1}(x)$ and so on]. It follows that the orbit of x is finite.

Conversely, let $x \in X$ be such that the orbit of x is finite. Then in the trajectory $x, f(x), f^2(x), \ldots$, there are only finitely many distinct terms. Let n be the least non-negative integer such that $f^n(x)$ repeats here. Then $f^n(x)$ is a periodic point and therefore x is an eventually periodic point. (And in fact periodic if $n = 0$). □

Theorem 1.10 *In every dynamical system (X, f) where X is a Hausdorff space,*

1. *$Fix(f)$ is a closed set.*
2. *$P(f)$ is an F_σ-set (that is, a countable union of closed sets).*
3. *$EP(f)$ is an F_σ-set.*
4. *$R(f)$ is a G_δ-set (that is, a countable intersection of open sets) if X is a metric space.*
5. *$\Omega(f)$ is a closed set.*

Proof 1. $Fix(f)$ is the set of all points where the continuous function f agrees with the identity map; therefore it is a closed set.

2. For each $n \in \mathbb{N}$, let $P_n(f) = \{x \in X \mid f^n(x) = x\} = Fix(f^n)$. Then by (1), each $P_n(f)$ is a closed set. We easily see that P is the union of these $P_n(f)$'s.

3. Here, we prove that $EP(f)$ is a F_σ-set.

$$\begin{aligned} EP(f) &= P(f) \cup f^{-1}(P(f)) \cup f^{-2}(P(f)) \cup \cdots \\ &= \text{a countable union of } F_\sigma\text{-sets} \\ &= \text{a } F_\sigma\text{-set.} \end{aligned}$$

4. Here, we prove that $R(f)$ is a G_δ-set.

$$\begin{aligned} R(f) &= \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \left\{x \in X \mid d(x, f^n(x)) < \frac{1}{k}\right\} \\ &= \text{a countable intersection of unions of open sets} \\ &= \text{a } G_\delta\text{-set.} \end{aligned}$$

5. Let $x \in \overline{\Omega(f)}$. We shall prove that x itself is non-wandering. For this, let V be a neighbourhood of x. Then V meets $\Omega(f)$. Take some $y \in V$ that is non-wandering. Because this V is a neighbourhood of that y, $\exists z \in V$ and $\exists n \in \mathbb{N}$ such that $f^n(z) \in V$. Since this is true for every neighbourhood V of x, it is a non-wandering point. □

Proposition 1.11 *$EP(f) \cap R(f) = P(f)$ holds in every dynamical system (X, f) if every finite set is closed in X.*

Proof We prove this when X is a metric space. We have already noted that every periodic point is both eventually periodic and recurrent (see Proposition 1.2 and the proof of statement (3) of Theorem 1.10). To prove the reverse inclusion, let x be both eventually periodic and recurrent. Let k be such that $f^k(x)$ is periodic. If x is not periodic, there is a positive distance δ from x to the finite orbit of $f^k(x)$. In the ball $B(x, \delta)$, only finitely many $f^i(x)$ lie. We can find a smaller ball $B(x, r)$ where no other term of the trajectory of x lies. So x is not recurrent. Thus, we have proved: If an eventually periodic point is not periodic, then it is not a recurrent point. □

Proposition 1.12 *Any two periodic orbits are disjoint or identical.*

Proof First we note that if x is a periodic point, then any two elements in its orbit, have the same orbit. i.e., the orbit of $f^m(x)$ is the same as the orbit of $f^n(x)$, even if $m \neq n$. Now if two periodic points x and y have a common point z in their orbits, the orbit of x = orbit of z and orbit of y = orbit of z and therefore orbit of x = orbit of y. □

Proposition 1.13 *If a trajectory converges, then its limit is a fixed point (in Hausdorff spaces).*

Proof Let $f^n(x) \to l$ as $n \to \infty$. Apply f. Because f is continuous, $f(f^n(x)) \to f(l)$. But $f(f^n(x))$ is a subsequence of $(f^n(x))$. Therefore it should converge to the same l. Thus $f(l) = l$. □

Here are some dynamical properties of subsets:

- Invariant set;
- Omega-limit set;
- Cycle.

Definition 1.14 Let (X, f) be a dynamical system.

1. A subset A of X is said to be invariant if $f(A) \subset A$. In that case $(A, f|_A)$ becomes a dynamical system It is called a subsystem of (X, f).
2. For $x \in X$, the omega-limit set of x, denoted by $\omega(f, x)$ is the set of all limit points of f-trajectory of x.
3. The orbit of a periodic point is called a cycle; the length of the cycle is the cardinality of the orbit.

Proposition 1.15 *In any dynamical system* (X, f)*, the sets* $Fix(f)$, $P(f)$, $EP(f)$, $R(f)$ *and* $\Omega(f)$ *are invariant sets.*

Proof 1. If x is a fixed point, then so is $f(x)$, because $f(x)$ is same as x.
2. If x is a periodic point, and if $f^n(x) = x$, then $f^n(f(x)) = f(f^n(x)) = f(x)$. So, $f(x)$ is a periodic point of the same period.
3. If x is eventually periodic, $\exists k \in \mathbb{N}$ such that $f^k(x) = y$ is periodic. Then by (2), $f^k(f(x)) = f(f^k(x)) = f(y)$ is also periodic. So $f(x)$ is also eventually periodic.
4. If x is a recurrent point, we prove that $f(x)$ is also recurrent. For this, let V be a neighbourhood of $f(x)$. Then (because f is continuous), $f^{-1}(V)$ is a neighbourhood of x. Because x is recurrent, $\exists n \in \mathbb{N}$ such that $f^n(x) \in f^{-1}(V)$. It implies that $f^{n+1}(x) \in V$. This is the same as $f^n(f(x)) \in V$. Thus $f(x)$ is also recurrent.
5. If x is a non-wandering point, we prove that $f(x)$ is also a non-wandering point. For this let V be a neighbourhood of $f(x)$. Then, $f^{-1}(V)$ is a neighbourhood of x. Since x is non-wandering $\exists y \in f^{-1}(V)$ and $\exists n \in \mathbb{N}$ such that $f^n(y) \in f^{-1}(V)$. This implies that $f(y) \in V$ and $f^n(f(y)) \in V$. This proves that $f(x)$ is non-wandering. □

2 Attracting Fixed Point

In this section, there are three subsections. In the first, a classical theorem about globally attracting fixed points is presented. In the second, various ways of understanding attracting fixed points are compared. In the third, several examples are provided that throw more light on the results proved in the first two sections.

2.1 Banach's Contraction Mapping Theorem

This is a theorem about globally attracting fixed point.

Definition 2.1 *(Contraction)* Let (X, d) be a metric space. Let $f : X \longrightarrow X$ be a self-map. We say that f is a contraction map if $\exists 0 < c < 1$ such that $d(f(x), f(y)) \leq c \cdot d(x, y)$ holds for all x, y in X. (Thus f contracts the distance between points).

Examples and non-examples: The map $f(x) = x^2$ is a contraction map on the interval $[-\frac{1}{4}, \frac{1}{4}]$ and the contraction constant c can be taken to be $\frac{1}{2}$. But it is not a contraction map on $[-1, 1]$ because $|0^2 - 1^2| = 1 = |0 - 1|$.

Proposition 2.2 *All contraction maps are uniformly continuous. In fact* $\forall x, y \in \mathbb{R}$ *and* $\forall \epsilon > 0$, $d(x, y) < \epsilon \implies d(f(x), f(y)) < \epsilon$.

Proposition 2.3 *A contraction map cannot have two fixed points.*

Proof If x, y are fixed points of f, then $d(f(x), f(y)) = d(x, y)$. So, the distance is not strictly reduced. □

Theorem 2.4 (Banach's Theorem) *Let (X, d) be a complete metric space (i.e., a space where all the Cauchy sequences are convergent). Let $f : X \longrightarrow X$ be a contraction map. Then f has a unique fixed point p. Moreover p is globally attracting (in the sense that every trajectory converges to p).*

Proof Let c be a contraction constant for f (i.e., let $0 < c < 1$ and $d(f(x), f(y)) \leq c \cdot d(x, y)$ holds for all points in X). Let $x \in X$. We first prove that its trajectory $(f^n(x))$ is a Cauchy sequence. As a first step, we claim:

$$d(f^n(x), f^{n+1}(x)) \leq c^n d(x, f(x)) \ \ \forall n \in \mathbb{N}.$$

This is proved by induction on n. When $n = 1$, this is from the definition of contraction map. If it is assumed for some $n \in \mathbb{N}$, then we can prove it for $n + 1$ as follows:

$$\begin{aligned} d(f^{n+1}(x), f^{n+2}(x)) &= d(f(s), f(t)) \qquad \text{where } s = f^n(x),\ t = f^{n+1}(x), \\ &\leq c \cdot d(s, t) \qquad \text{because } c \text{ is contraction constant} \\ &= c \cdot d(f^n(x), f^{n+1}(x)) \qquad \text{by re-substitution} \\ &\leq c \cdot c^n \cdot d(x, f(x)) \qquad \text{by induction hypothesis} \\ &= c^{n+1} \cdot d(x, f(x)). \end{aligned}$$

Therefore, the principle of induction completes the proof of our claim. For us, this is not the goal but is merely an intermediate step. Now let $m < n$. Then

$$\begin{aligned} d(f^m(x), f^n(x)) &\leq d(f^m(x), f^{m+1}(x)) + d(f^{m+1}(x), f^{m+2}(x)) + \cdots \\ &\quad + \cdots + d(f^{n-1}(x), f^n(x)) \qquad \text{(by triangle inequality)} \\ &\leq c^m \cdot d(x, f(x)) + c^{m+1} \cdot d(x, f(x)) + \cdots \\ &\quad + \cdots + c^{n-1} \cdot d(x, f(x)) \qquad \text{(by using our claim)} \\ &= d(x, f(x))(c^m + c^{m+1} + \cdots + c^{n-1}). \end{aligned}$$

We already know that the geometric series $\Sigma_{n=0}^{\infty} c^n$ converges (here, we use the fact that $0 < c < 1$) and therefore by the Cauchy principle of convergence, for every $\epsilon > 0$ there is some $n_0 \in \mathbb{N}$ such that for $n > m \geq n_0$ the finite sum $c^m + c^{m+1} + \cdots + c^{n-1} < \epsilon$. It follows that $f^m(x)$ is a Cauchy sequence because if m and n are $\geq n_0$, we have

$$d(f^m(x), f^n(x)) \leq \epsilon \cdot d(x, f(x)),$$

and here $d(x, f(x))$ is a constant (not depending on m or n).

Next, because (X, d) is assumed to be complete, this Cauchy sequence should converge to some point p in X. By Proposition 1.13, this p should be a fixed point. Thus, we have proved that the unique fixed point p is globally attracting. □

Remark 2.5 If (X, d) is any metric space (not necessarily complete) and if f is a contraction map on it, having a fixed point p, we can prove that p is globally attracting. This is because $\forall x \in X$ and $n \in \mathbb{N}$, we have

$$d(p, f^n(x)) \leq c^n \cdot d(p, f(x)),$$

as can be proved by induction. It follows that $(f^n(x))$ converges to p.

2.2 Various Versions of Attraction

Theorem 2.6 *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be continuously differentiable and let $f(p) = p$. Then each statement below implies the next:*

1. $|f'(p)| < 1$.
2. *f is a local contraction at p.*
3. $\bigcap_{n=1}^{\infty} f^n(V) = \{p\}$ *for some f-shrinking neighbourhood V of p (in the sense $f(\overline{V}) \subset V$).*
4. $\{x \in \mathbb{R} \mid f^n(x) \to p\}$ *has the point p in its interior.*
5. $|f'(p)| \leq 1$.

Proof (1) $\Longrightarrow$ (2) : Take r such that $|f'(p)| < r < 1$. Because f' is continuous, there is $\delta > 0$ such that $|f'(x)| < r$ holds for all $x \in (p - \delta, p + \delta)$. We now prove that $J = [p - \delta, p + \delta]$ is f-invariant and that $f|_J$ is a contraction map.
If $x \in J$, then $|\frac{f(x)-f(p)}{x-p}| = |f'(c)|$ for some c between x and p (by the mean value theorem) and is therefore $< r$. In particular, $|f(x) - f(p)| < |x - p|$ and this proves (because $f(p) = p$) that $f(x)$ is nearer to p than x; so that $f(x) \in J$. Also if $x \neq y \in J$, then $|\frac{f(x)-f(y)}{x-y}| = |f'(c)|$ for some c between x and y, and is therefore $< r$. This means that f is a contraction map on J with r as the contraction constant.

(2) $\Longrightarrow$ (3) : Here, we do not assume the differentiability of f. We prove more generally that in any locally compact metric space X, (2) implies (3).

Let f be a local contraction at $p = f(p)$ in the sense that there is a neighbourhood V of p such that V is f-invariant and such that the restriction of f to V is a contraction map. We first note that any ball $B(p, r)$ that is $\subset V$ will be f-invariant. (Because if $x \in B(p, r)$, then $d(f(x), f(p)) \leq c \cdot d(x, p)$ where c is the contraction constant; this gives that $f(x)$ is nearer to p than x). Because X is locally compact, one such ball $B(p, r)$ has compact closure. Call it W. We note that W is f-shrinking in the sense that $f(\overline{W}) \subset W$. In fact, if $x \in \overline{W}$, then $d(f(x), f(p) = p) \leq c \cdot d(x, p) < c \cdot r$ and so $f(x) \in W$.

Next, we claim that $\bigcap_{n=1}^{\infty} f^n(W)$ is $\{p\}$. In fact, $f(W) \subset B(p, cr)$ and repeated use of the previous argument gives $f^n(W) \subset B(p, c^n r)$ for every positive integer n. Therefore, $\bigcap_{n=1}^{\infty} f^n(W) \subset \bigcap_{n=1}^{\infty} B(p, c^n r) = \{p\}$ (since $c^n \to 0$).

(3) $\Longrightarrow$ (4) : Here, we do not assume that the domain X is a metric space. It can be any locally compact Hausdorff space.

Let V be any f-shrinking neighbourhood of p, as above. We claim that for every $x \in V$, the trajectory $(f^n(x))$ converges to p. Consider the decreasing sequence of compact subsets $\left\{f^n(\overline{V})\right\}$ (because $f^{n+1}(\overline{V}) \subset f^n(f(\overline{V}))$). We have $\bigcap_{n=1}^{\infty} f^n(\overline{V})$ is the same as $\bigcap_{n=1}^{\infty} f^n(V)$ (because $f^{n+1}(\overline{V}) = f^n(f(\overline{V})) \subset f^n(V)$ as V is f-shrinking and $f^n(V) \subset f^n(\overline{V})$ as $V \subset \overline{V}$) and that is given to be $\{p\}$. If W is any neighborhood of p, it will swallow $f^n(\overline{V})$ for some n (because of compactness; otherwise $(f^n(\overline{V}) - W)$ will be a decreasing sequence of compact non-empty sets with empty intersection, leading to a contradiction).

Now if $x \in V$, the sequence $f^n(x)$ has to be eventually in W. This is true for every neighbourhood W of p. This means that $f^n(x) \to p$.

(4) $\Longrightarrow$ (5) : Here, of course we assume that f is continuously differentiable. Let if possible $|f'(p)| > 1$. Choose s such hat $|f'(p)| > s > 1$ and choose a ball $B(p, r)$ such that $|f'(x)| > s$ for all x in that ball. If $x \in B(p, r)$, then we have $|f(x) - f(p)| = |f'(c)| \cdot |x - p|$ for some c in that ball (by mean value theorem) and this is $> s \cdot |x - p|$.

If $f(x)$ is also in the ball, then $f(f(x))$ is further away from p than $f(x)$. Since $s^n \to \infty$, the sequence $f^n(x)$ leaves the ball at some time or other. Even if it enters the ball at a later time, by the above argument, unless it hits p precisely it has to leave the ball again. This can be used to contradict (4), (see Exercise (9)).

We have actually proved more. No element in the ball gets attracted to p. (We started to prove the weaker statement that some element doesn't get attracted to p, but proved that p is repelling as per the Definition 2.11). □

Remark 2.7 Note that the difference between (1) and (5) is slight. But it is worthwhile to point out that (2) implies (1). In fact $|f'(p)|$ can not exceed the contraction constant. Thus (1) and (2) provide the strong form of attraction.

2.3 Examples

Example 2.8 Consider $f : \left[-\frac{1}{2}, 1\right] \longrightarrow [0, 1]$ defined by

$$f(x) = \begin{cases} -2x & \text{if } x \leq 0 \\ \frac{x}{2} & \text{if } x \geq 0. \end{cases}$$

Illustration: Then, f has a unique fixed point 0; it is globally attracting ; but f is not a contraction map; f is not differentiable at 0. □

Example 2.9 Define $f : [0, 1] \longrightarrow [0, 1]$ by specifying $f(\frac{1}{n}) = \frac{1}{n+1}$, $\forall n \in \mathbb{N}$ and f is linear on each $[\frac{1}{n+1}, \frac{1}{n}]$.

Illustration: Then, the range of f is $[0, \frac{1}{2}]$. Still, f is not a contraction because; if c were the contraction constant,

$$\left| f\left(\frac{1}{n}\right) - f(0) \right| = \frac{1}{n+1} - 0 \leq c \cdot \frac{1}{n} \implies c \geq \frac{n}{n+1}, \ \forall n \in \mathbb{N}.$$

Therefore, $c \geq 1$, contrary to the definition. But, f has a unique fixed point (namely 0) and it is globally attracting as well. Every trajectory is converging to 0. This shows that the converse of the contraction mapping theorem (i.e, Theorem 2.4) is not true. (This also shows that in Theorem 2.6, (3) does not imply (2)). □

Example 2.10 Let $f(x) = x - x^3$ on $\mathbb{R}$.

Illustration: Then, the only fixed point of f is 0. This is because $x - x^3 = x$ has only one solution. Even though $f'(0) = 1$ (and therefore 0 is not attracting as per the calculus definition), we shall now prove that its basin of attraction contains the closed interval $[-1, 1]$ (and it is seen to be a neighbourhood of 0).

If $0 < x < 1$, then $x^3 < x$ and therefore $0 < x - x^3 = f(x) < x < 1$. Repeatedly using this, we find that the trajectory $(f^n(x))$ is strictly decreasing and bounded. It has to converge (and by proposition of the last section) its limit should be a fixed point. It has to be 0 because there is no other fixed point. Similarly, if $-1 < x < 0$, then the trajectory of x is strictly increasing and converges to 0. 0 is an attracting fixed point in all other senses except (1) of Theorem 2.6. The points 1 and -1 even map to the fixed point 0. □

Definition 2.11 1. **Repelling Fixed Point:** In a dynamical system (X, f), a fixed point x is said to be repelling if $\exists$ a neighbourhood V of x such that $\forall y \neq x$ in V, $\exists n \in \mathbb{N}$ such that $f^n(y) \notin V$.
2. **Attracting Cycle:** In a dynamical system (X, f), a cycle of length n is said to be attracting if every element in it, is an attracting fixed point of f^n.

Example 2.12 Consider the map $f(x) = x^2 - 4x + 5$.

Illustration: There are two fixed points $\dfrac{5+\sqrt{5}}{2}$ and $\dfrac{5-\sqrt{5}}{2}$. Both are repelling. Because $|f'|$ at these points is calculated as $\sqrt{5} + 1$ and $\sqrt{5} - 1$, and both these are > 1.

Next, we calculate that 1 and 2 are periodic points of period 2. Observe that

$$f(1) = 1^2 - 4 + 5 = 2 \quad \text{and} \quad f(2) = 2^2 - 8 + 5 = 1.$$

Now we shall prove that this 2-cycle $\{1, 2\}$ is attracting.

$$|(f \circ f)'(1)| \;=\; |f'(f(1))f'(1)| \;=\; |f'(1)||f'(2)| \;=\; |2-4||4-4| \;=\; 0.$$

To understand the nature of attraction without calculus, we take the point $\frac{3}{2}$. Using a calculator, we find that its trajectory is

$$\frac{3}{2}, \frac{5}{4}, \frac{25}{16}, \frac{305}{256}, \frac{108385}{65536}, \ldots$$

and the same becomes in the decimal notation

$$1.5, 1.25, 1.5625, 1.191, 1.6538, 1.11984, 1.11984, 1.77468, 1.05077 \cdots .$$

We notice the following in this sequence. If we take the even terms

$$1.25, 1.191, 1.11984, 1.05077, \ldots$$

it is strictly decreasing and going nearer and nearer to the periodic point 1. If we take the remaining terms

$$1.5, 1.5625, 1.6538, 1.77468, \ldots$$

it is strictly increasing, going nearer and nearer to the other periodic point 2. We guess that the same happens to all points in a neighbourhood of 1 and 2 (Prove it!). □

Example 2.13 Let $f(x) = x + \sin x$ on $[0, 4\pi]$.

Illustration: The fixed points of f are easily calculated as $0,\ \pi,\ 2\pi,\ 3\pi$ and 4π. Which of them are attracting and which of them are repelling? To see this, we calculate

$$f'(x) = 1 + \cos x; \quad f'(0) = 2; \quad f'(\pi) = 0; \quad f'(2\pi) = 2; \quad f'(3\pi) = 0; \quad f'(4\pi) = 2.$$

We conclude that π and 3π are attracting points, whereas the other three are repelling. □

Among the examples seen here, this is the only interval map with two attracting fixed points. Its analogues can be easily constructed to show that $\forall n \in \mathbb{N}$, there is an interval map with n attracting fixed points.

This example serves another purpose also. This f is a homeomorphism from $[0, 4\pi]$ to itself. Its graph seems to be winding around the diagonal. If p is a fixed point, see

whether in the neighbourhood of (p, p), the graph is above or below the diagonal (until the adjacent fixed point). Then p is attracting if and only if the graph is above the diagonal to the left of p, and below the diagonal to the right of p. This seems to suggest the following statement.

Let f be an one-one interval map with a discrete set of fixed points. Then between any two attracting fixed points, there has to be a repelling fixed point and vice-versa. (see Exercise 8.13). □

Example 2.14 Let $f(x) = 2x - 2x^2$ on I. This is one of the logistic maps introduced earlier.

Illustration: Here, f has two fixed points, namely the two roots of $2x - 2x^2 = x$. These are 0 and $\frac{1}{2}$. We have $f'(x) = 2 - 4x$, $f'(0) = 2$ and $f'(\frac{1}{2}) = 0$. Therefore 0 is a repelling fixed point and $\frac{1}{2}$ is an attracting fixed point.
We find the basin of attraction for $\frac{1}{2}$. If p is a fixed point, then its basin of attraction is defined as the set of all points whose trajectories converge to p. First, from the equation $2x - 2x^2 = -2(x - \frac{1}{2})^2 + \frac{1}{2}$, we find that $\frac{1}{2}$ is the maximum value. We find that the graph is above the diagonal on $\left[0, \frac{1}{2}\right)$ and below the diagonal on $\left(\frac{1}{2}, 1\right]$. This makes us guess

$$x < 2x - 2x^2 < \frac{1}{2} \text{ if } 0 < x < \frac{1}{2} \text{ and}$$
$$x > 2x - 2x^2 \text{ if } \frac{1}{2} < x < 1.$$

We in fact have $0 < x < \frac{1}{2} \Longrightarrow 0 < 2x < 1 \Longrightarrow 0 < 2x^2 < x \Longrightarrow x < 2x - 2x^2$. And $\frac{1}{2} < x < 1 \Longrightarrow x > 2x - 2x^2$ because $1 > 2 - 2x$.

For all $0 < x < \frac{1}{2}$, the trajectory of x increases and remains in $\left(0, \frac{1}{2}\right]$. So, it has to converge to a fixed point. The only available fixed point is $\frac{1}{2}$. Points in $\left(\frac{1}{2}, 1\right)$, at the very next time fall in $\left(0, \frac{1}{2}\right)$ and afterwards go nearer and nearer to $\frac{1}{2}$. The basin of attraction is $(0, 1)$. □

3 Topological Transitivity

In this section, an important dynamical property known as topological transitivity, is studied. In the first subsection, it is understood in five different (but equivalent) ways. In the second, five different methods to prove topological transitivity are introduced.

Notation: We will use the term **opene set** as our abbreviation for an **open nonempty set**.

3.1 Five Views of Topological Transitivity

There are five ways in which topological transitivity may be understood. To put them roughly:

1. It is possible to go from any sub-region to any other sub-region (as the name 'transitive' suggests) (High mobility).
2. Every sub-region is visited by plenty of points at some time or other (Abundance of visitors).
3. Every invariant set is either too small or too big (Lack of invariant sets).
4. There is no nontrivial proper subsystem (Triviality of sub-system/ Indecomposability).
5. There is a dense orbit (Highly wandering points).

These five are precisely stated as follows: Let (X, f) be a dynamical system.

1. If V and W are any two nonempty open subsets of X, there is $n \in \mathbb{N}$ such that $f^n(V) \cap W$ is nonempty.
2. If V is a nonempty open set, then $\{x \in X \mid f^n(x) \in V \text{ for some } n \in \mathbb{N}\}$ is dense.
3. If $A \subset X$ is such that $f(A) \subset A$, then either A is dense or A is nowhere dense.
4. $(Y, f|_Y)$ is called a closed subsystem of (X, f) if Y is a closed subset of X and $f(Y) \subset Y$ so that $f|_Y$ is a self-map of Y. The present requirement is: if $(Y, f|_Y)$ is a closed subsystem of (X, f), then either $Y = X$ or Y has empty interior.
5. $\exists x \in X$ such that $\overline{O(x, f)} = X$.

Theorem 3.1 *If X is a compact metric space without isolated points (like the unit interval I, the circle $\mathbb{S}^1$, the torus $\mathbb{T}$, the space Σ of sequences, etc.), then the above five are equivalent.*

Proof (1) $\Longrightarrow$ (2) : If W is an opene set, then

$$\left\{x \in X \mid f^n(x) \in W \text{ for some } n \in \mathbb{N}_0\right\}$$

can be rewritten as $\bigcup_{n \in \mathbb{N}_0} f^{-n}(W)$. Let us denote this set by W^*. To prove that W^* is dense in X, take any opene set V in X. Then by (1), there is $n \in \mathbb{N}$ such that $f^n(V) \cap W$ is nonempty. This implies that $f^{-n}(W) \cap V$ is nonempty and therefore $W^* \cap V$ is nonempty. Since this is true for every opene set V, it follows that W^* is dense.

(2) $\Longrightarrow$ (3) : Let $A \subset X$ be an invariant set so that $f(A) \subset A$. Then its complement has the property that no element of A visits it at any time. This means $(A^c)^* = A^c$ in the notation of the previous paragraph. Now consider two cases.

1. If A^c has an interior point, then (2) implies that it is dense. In this case A has empty interior.
2. In the other case A^c has empty interior.

So we have that either A or A^c has empty interior. Equivalently, either A or A^c is dense. Now, $\overline{A}$ is also f-invariant (follows from the continuity of f). Therefore, if A is not dense, then $\overline{A}$ has no interior and is thus, nowhere dense.

(3) $\Longrightarrow$ (4) : Let $(Y, f|_Y)$ be a closed subsystem of (X, f). Then Y is f-invariant subset of X. Therefore by (3), Y or Y^c is dense. But Y is closed. Therefore either $Y = X$ or Y is nowhere dense.

(4) $\Longrightarrow$ (2) : Let W be a nonempty open set. Look at $(W^*)^c$. This is the set of all points that never visit W. It can be written as $\left(\bigcup_{n\in\mathbb{N}_0} f^{-n}(W)\right)^c$. Because f is continuous and W is open, the subset $f^{-n}(W)$ is open for each $n \in \mathbb{N}$. Therefore W^* is open and $(W^*)^c$ is closed. Also it is f-invariant. This is because if x never visits W, so is the case with $f(x)$, [for if $f(x)$ visits W at time n, then x will visit W at time $n + 1$]. Thus $(W^*)^c$ gives rise to a sub-system of (X, f). By (4), either $(W^*)^c = X$ or $(W^*)^c$ has empty interior. The former case is not possible because W is nonempty. In the latter case W^* is dense. This proves (2).

(2) $\Longrightarrow$ (5) : Here we use the Baire Category theorem. Because we have assumed that X is a compact metric space, X admits a countable base say $\{B_1, B_2, B_3, \ldots\}$ of nonempty open sets. By (2), each B_n^* is dense and open. Therefore by the Baire Category theorem, $\bigcap_{n\in\mathbb{N}} B_n^*$ is non empty. If x is a point in this intersection, then x visits each B_n at some time or other. In other words, the orbit of x meets every basic open set and is therefore dense.

(5) $\Longrightarrow$ (1) : Let the orbit of x be dense. Let V, W be two opene sets. First choose $n_1 \in \mathbb{N}$ such that $f^{n_1}(x) \in V$. Next take the set

$$W - \left\{x, f(x), \ldots, f^{n_1}(x)\right\}.$$

This is open (because every finite set is closed) and non-empty (because W is infinite, because there are no isolated points). The orbit of x should meet this opene set also. Therefore $\exists n_2 > n_1$ in $\mathbb{N}$ such that $f^{n_2}(x)$ meets W. Thus $f^{n_2-n_1}(f^{n_1}(x)) \in W$ and so $f^{n_2-n_1}(V)$ meets W. □

3.2 Five Proofs of Topological Transitivity of the Tent Map

Recall that the tent map $f : I \longrightarrow I$ is defined by the formula

$$f(x) \;=\; \begin{cases} 2x & \text{if } 0 \le x \le \frac{1}{2} \\ 2 - 2x & \text{if } \frac{1}{2} \le x \le 1. \end{cases}$$

Equivalently, this is given by the formula $f(x) = 1 - |1 - 2x|$ for all x. It is the piece-wise linear map specified by

$$f(0) = 0, \quad f\left(\frac{1}{2}\right) = 1, \quad f(1) = 0.$$

Theorem 3.2 *The tent map is topologically transitive.*

Proof The reason for providing five different proofs is that each proof leads to a more general theorem; at the end we will have five theorems each giving a set of assumptions that implies topological transitivity. □

First Proof: Call a map locally eventually onto (in short, **l.e.o.**) if for every opene set V, $\exists n \in \mathbb{N}$ such that $f^n(V) = I$. We shall prove the following two assertions:

1. Every l.e.o. map is topologically transitive.
2. The tent map is l.e.o.

In the literature, l.e.o maps are sometimes known as **strongly transitive** maps and also known as **topologically exact** maps.

Proof of (1): Let V and W be two opene sets. By assumption $\exists n \in \mathbb{N}$ such that $f^n(V) = I$. Then obviously $f^n(V) \cap W$ is nonempty. This proves that f is topologically transitive.

Proof of (2): This is divided into two parts:

(2a) Every opene set must contain a dyadic sub-interval.
(2b) Every dyadic sub-interval J admits $n \in \mathbb{N}$ such that $f^n(J) = I$.

Here, a dyadic sub-interval is a closed interval of the form $I_{k,m} = \left[\frac{k-1}{2^m}, \frac{k}{2^m}\right]$ where $1 \leq k \leq 2^m$, k and m are non-negative integers. It is easily noted that the length of $I_{k,m}$ is $\frac{1}{2^m}$. The union of $I_{k,m}$ for a fixed m, as k varies, is I. There are 2^m sub-intervals of the form $I_{k,m}$ (as k varies).

$$I_{k_1,m_1} \subset I_{k_2,m_2} \implies m_1 \geq m_2.$$

Proof of (2a): Let V be an opene set. Then V contains a closed interval J of positive length l. Choose a positive integer n large enough so that $\frac{1}{2^n} < \frac{l}{2}$. Let k be the least non-negative integer such that $\frac{k}{2^n} \in J$. Then we are sure that $\frac{k+1}{2^n}$ also belongs to J. This implies that the dyadic sub-interval $I_{k,n} \subset J \subset V$.

Proof of (2b): We prove by induction on m that $f^m(I_{k,m}) = I$. When $m = 1$, we have two such sub-intervals $I_{0,1}$ and $I_{1,1}$ of length $\frac{1}{2}$. We directly verify that $f(I_{0,1}) = I$ and $f(I_{1,1}) = I$. Suppose by induction hypothesis, we have proved that

$f^m(I_{k,m}) = I$ for some $m \in \mathbb{N}$ and for all $k,\ 0 \le k \le 2^m$. Then we prove a similar result for $m+1$. For this, we first note that $f(I_{k,m+1}) = I_{r,m}$ for some r. Then using induction hypothesis,

$$f^{m+1}(I_{k,m+1}) \ = \ f^m(f(I_{k,m+1})) \ = \ f^m(I_{r,m}) \ = \ I.$$

The principle of induction now completes the poof of (2b).

Second Proof :

Definition 3.3 An interval map is said to be length-expanding if there is $\delta > 0$ such that the following holds:
$|f(J)| \ge (1+\delta)|J|$ for all intervals J of positive length unless $f(J) = I$.

Note that this implies $f(J) = I$ for all sub-intervals J of length $> \frac{1}{1+\delta}$. In particular, an interval map is length-doubling if $|f(J)| \ge 2|J|$ holds unless $f(J) = I$.

We split this proof into three parts:

(A) Every length-expanding interval map is topologically transitive.
(B) If f is the tent map, $f \circ f$ is length-doubling (and therefore length-expanding).
(C) If $f \circ f$ is topologically transitive, so is f.

Proof of (A): We shall prove that f is l.e.o. Let V be an opene set. Then V contains a closed interval J of positive length. By our assumption, $|f(J)| \ge (1+\delta)|J|$ unless $f(J) = I$. This $f(J)$ is also an interval of positive length, and so $|f^2(J)| \ge (1+\delta)^2|J|$ unless $f^2(J) = I$.

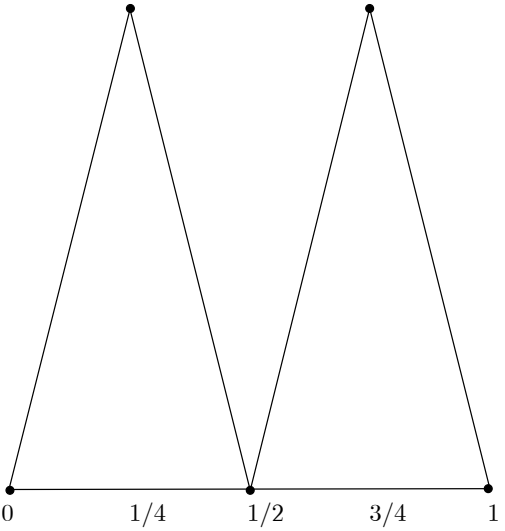
Repeating this argument, we obtain that for every $n \in \mathbb{N}$,

$$|f^n(J)| \ \ge \ (1+\delta)^n|J|,$$

unless $f^n(J) = I$. Because the sequence $(1+\delta)^n$ is growing exponentially, there is $n \in \mathbb{N}$ such that $(1+\delta)^n|J| > |I|$. For that n, we do have $f^n(J) = I$. It follows that $f^n(V) = I$.

Proof of (B): The graph of $f \circ f (= g)$ is as sketched below. It is a piece-wise linear map specified by

$$g(0) = 0,\ g\left(\frac{1}{4}\right) = 1,\ g\left(\frac{1}{2}\right) = 0,\ g\left(\frac{3}{4}\right) = 1,\ \text{and } g(1) = 0.$$



On $\left[0, \frac{1}{4}\right]$, this function is $f \circ f(x) = 4x$. Therefore, if J is an interval and is a subset of $\left[0, \frac{1}{4}\right]$, then $|g(J)| = 4|J|$. Then, similar statement holds depending on $J \subset \left[\frac{1}{4}, \frac{1}{2}\right]$, $J \subset \left[\frac{1}{2}, \frac{3}{4}\right]$ or $J \subset \left[\frac{3}{4}, 1\right]$.

If J is not contained in any of these four, we argue as follows: Suppose $J \subset \left[0, \frac{1}{2}\right]$. Then $J = J_1 \cup J_2$ where $J_1 = J \cap \left[0, \frac{1}{4}\right]$ and $J_2 = J \cap \left[\frac{1}{4}, \frac{1}{2}\right]$. By previous paragraph

$$|g(J_1)| = 4|J_1| \quad \text{and} \quad |g(J_2)| = 4|J_2|.$$

Being aware that $g(J_1)$ and $g(J_2)$ may overlap, we deduce that

$$|g(J)| \geq \text{ both } 4|J_1| \text{ and } 4|J_2|$$

and moreover, $|g(J)| \geq 2|J|$. In the last step, we make use of the fact that $|J_1|$ or $|J_2|$ should be at least $\frac{1}{2}|J|$. This proves that $|g(J)| \geq 2|J|$ for all sub-intervals J that are subsets of $\left[0, \frac{1}{2}\right]$.

Analogous statements hold for sub-intervals $\subset \left[\frac{1}{4}, \frac{3}{4}\right]$, and for sub-intervals $\subset \left[\frac{1}{2}, 1\right]$.

The only remaining case is when J is not contained in any of these three. Then we can prove that J should contain one of the four sub-intervals $\left[0, \frac{1}{4}\right]$, $\left[\frac{1}{4}, \frac{1}{2}\right]$, $\left[\frac{1}{2}, \frac{3}{4}\right]$, $\left[\frac{3}{4}, 1\right]$. Then $f^2(J) = I$. We have thus completed the proof that $f \circ f$ is length-doubling.

Remark 3.4 Observe that the tent map is not length expanding as

$$f\left(\left[\frac{1}{4}, \frac{3}{4}\right]\right) = \left[\frac{1}{2}, 1\right].$$

Proof of (C): More generally, let (X, f) be a dynamical system and let $(X, f \circ f)$ be topologically transitive. To prove that f is also so, let V and W be two opene sets. By assumption there is $n \in \mathbb{N}$ such that $(f \circ f)^n(V)$ meets W. It follows that $f^{2n}(V)$ meets W. This proves f is also topologically transitive.

Third proof: We directly exhibit a point in I whose orbit under the tent map f is dense in I. Let s be a Morse sequence over $\{0, 1\}$, that is, an infinite sequence of 0's and 1's such that every word w over $\{0, 1\}$ occurs in that sequence. If $s = s_1 s_2 s_3 \cdots s_n \cdots$ then form the sets $A_n = \{x \in I \mid f^n(x) \in I_{s_n}\}$ for all $n \in \mathbb{N}$, where $I_0 = \left[0, \frac{1}{2}\right]$ and $I_1 = \left[\frac{1}{2}, 1\right]$.

We note that $\bigcap_{n\in\mathbb{N}} A_n$ is nonempty, by Cantor-Intersection theorem. This is because the finite intersections $A_1 \cap A_2 \cap \cdots \cap A_n$ are nonempty and closed for each $n \in \mathbb{N}$. Actually, this is an interval of length $\frac{1}{2^n}$.

Let x be a point common to all these intervals. Then we now prove that the orbit of x is dense. For this the following notation will be convenient. The dyadic sub-intervals defined in the second proof are now provided with a different notation:

$$I_0 = \left[0, \frac{1}{2}\right] \quad I_1 = \left[\frac{1}{2}, 1\right];$$

$$I_{00} = \left[0, \frac{1}{4}\right] \quad I_{01} = \left[\frac{1}{4}, \frac{1}{2}\right] \quad I_{10} = \left[\frac{3}{4}, 1\right] \quad I_{11} = \left[\frac{1}{2}, \frac{3}{4}\right];$$

$$I_{000} = \left[0, \frac{1}{8}\right] \quad I_{001} = \left[\frac{1}{8}, \frac{1}{4}\right] \quad \text{and so on.}$$

In general, if w is a word $w_1 w_2 w_3 \cdots w_n$ over $\{0, 1\}$, then I_{w0} is defined as $I_w \cap f^{-k}(I_0)$; I_{w1} is defined as $I_w \cap f^{-k}(I_1)$, where $k = \mid w \mid$. Recursively, this defines I_w for every word w. We see that these $I'_w s$ are same as the dyadic sub-intervals. In particular every interval J of positive length must contain one of these I_w's.

Now to prove that the orbit of x is dense in I, let J be an open sub-interval of positive length. Then J contains I_w for some word w. This w occurs in the Morse sequence s. Therefore there is $n \in \mathbb{N}$ such that $w = s_n s_{n+1} \cdots s_{n+r}$.

Now the point $f^n(x)$ has the property that it belongs to I_w. (Because $f^n(x) \in I_{s_n}$, $f^{n+1}(x) \in I_{s_{n+1}}$, etc). Here we use the fact that

$$I_w = I_{w_1 w_2 \cdots w_n} = \{x \in I \mid x \in I_{w_1}, f(x) \in I_{w_2}, f^2(x) \in I_{w_3}, \ldots\}.$$

Thus $f^n(x) \in J$ for this n.

Fourth Proof: The tent map is an example of a Markov map. We consider more generally, the interval maps f for which there is a finite set of points $0 = x_0 < x_1 < \cdots < x_n = 1$ such that if $I_i = \left[x_i, x_{i+1}\right]$, then the following are satisfied:

(i) $|f'(x)| > 1$ for all x except those finitely many points where f may not be differentiable.
(ii) For each i, $f(I_i)$ is the union of some I_j's. (This means : If $f(I_i)$ meets some I_j, then it contains it).
(iii) f is monotonic on each I_i.

We shall call such maps **EMI maps** (Expanding Markov Interval map). The Markov matrix associated with such a map f is a $n \times n$ matrix where the (i, j)- entry is 1 if $f(I_i) \supset I_j$, and (i, j)- entry is 0 otherwise.

We say that a matrix of 0's and 1's is irreducible if for every (i, j), there is k in $\mathbb{N}$ such that the (i, j)-entry of the kth power of that matrix is nonzero.Our fourth proof of transitivity of the tent map is as usual divided into two parts, one giving a general theorem and the other stating that the tent map satisfies its assumptions.

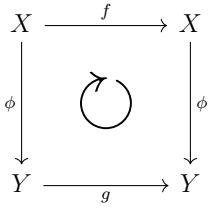
(A) Every EMI map whose matrix is irreducible, is topologically transitive.
(B) The tent map is an EMI map with irreducible matrix.

Proof of (A): For the proof, see Proposition (4.8) in [14].

Proof of (B): Take $0 = x_0 < x_1 = \frac{1}{2} < x_2 = 1$, $I_0 = \left[0, \frac{1}{2}\right]$, $I_1 = \left[\frac{1}{2}, 1\right]$. Then $f(I_0) = I_0 \cup I_1 = I$; $f(I_1) = I_0 \cup I_1 = I$ and $|f'(x)| = 2$ for all x other than x_i's. Thus we have easily verified that the tent map is an EMI map. Its Markov matrix is $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and is obviously irreducible.

Fifth Proof:

Definition 3.5 *(Topological Conjugacy:)* If (X, f) and (Y, g) are two dynamical systems, a homeomorphism $\phi : X \longrightarrow Y$ is called a topological conjugacy if the following diagram commutes:



Remark 3.6 All dynamical properties, seen so far, are preserved by topological conjugacy.

Here the main idea is that the well known shift map (Σ_2, σ) and this tent map (I, f) (though they are not topologically conjugate themselves) admit dense subsystems (D, σ) and (D_2, f) respectively, that are topologically conjugate to each other. We shall then use the following lemma.

Lemma 3.7 *Let (X, f) be a dynamical system and Y be a dense subset of X such that $f(Y) \subset Y$. Then (X, f) is topologically transitive if and only if $(Y, f|_Y)$ is topologically transitive.*

Proof Let (X, f) be topologically transitive. To prove that $(Y, f|_Y)$ is also topologically transitive, let V and W be two opene sets in Y. Then there are open sets $\widetilde{V}$ and $\widetilde{W}$ in X such that $V = \widetilde{V} \cap X$ and $W = \widetilde{W} \cap Y$. Because f is topologically transitive, $\exists x \in \widetilde{V}$ and $n \in \mathbb{N}$ such that $f^n(x) \in \widetilde{W}$. But alas! this x may not be in Y. By the continuity of f^n, there is a neighborhood U of $x \in X$ (we may assume that $U \subset \widetilde{V}$) such that $f^n(U) \subset \widetilde{W}$. Now because Y is dense in X, this U meets Y. Take y in $U \cap Y$. Then, $f^n(y) \in \widetilde{W}$. Also, because Y is invariant, $f^n(y) \in Y$ also. Thus $f^n(y) \cap W$ is nonempty.

Conversely, let $(Y, f|_Y)$ be topologically transitive. To prove that (X, f) is topologically transitive let V and W be two opene sets in X. Then $V \cap Y$ and $W \cap Y$ are opene sets in Y. (Here we use the fact that Y is dense in X). Because $f|_Y$ is topologically transitive, $\exists n \in \mathbb{N}$ such that $f^n(V \cap Y) \cap (W \cap Y)$ is nonempty. It follows that $f^n(V) \cap W$ is nonempty. □

Remark 3.8 It is interesting to note that the subsystem Y, although topologically transitive need not contain a dense orbit. For example, for the tent map if $Y = \mathbb{Q} \cap I$, then all $y \in Y$ are pre-periodic, hence no dense orbit, but here Y is countable and not a Baire space. A Baire space is one in which countable intersection of open and dense sets is again dense.

Now we proceed to apply this lemma in the present context. We take

$$\begin{aligned} D_1 &= \{x \in I \mid f^n(x) \neq 0 \text{ for all } n \in \mathbb{N}\} \\ &= \text{complement of the backward obit of 0.} \end{aligned}$$

We can prove that $f^n(x) = 0$ for some $n \in \mathbb{N}$ if and only if x is a dyadic rational number. That is, $D_1 =$ complement in I of dyadic rational numbers (a rational number whose denominator is a power of 2). Since the complement of a countable set is always dense, it follows that D_1 is a dense subset of I.

If $x \in D_1$, then $f(x) \in D_1$, because if $f^n(f(x)) = 0$, then $f^{n+1}(x) = 0$. We consider the subsystem $(D_1, f|_{D_1})$. On the other side, we start with the shift map (Σ_2, σ). Let $D_2 = \{x \in \Sigma_2 \mid x \text{ is not eventually0}\}$. Then D_2 is the complement of a countable set and is therefore dense in Σ_2 (because in Σ_2 also, every opene set is uncountable). Therefore by the above lemma, $(D_2, \sigma|_{D_2})$ is topologically transitive.

We now construct a homeomorphism $\phi : D_1 \longrightarrow D_2$ as follows: If $x \in D_1$, then

$$(\phi(x))_n = \begin{cases} 0 & \text{if } f^n(x) < \frac{1}{2}; \\ 1 & \text{if } f^n(x) > \frac{1}{2}. \end{cases}$$

(Note that $f^n(x)$ cannot be equal to $\frac{1}{2}$ if $x \in D_1$). Note that $\phi(x) \in D_2$.

Lemma 3.9 *The following are equivalent for x and y in D_1 and for $n \in \mathbb{N}$.*

(i) There is a word w of length n such that x and y belong to the same I_w.

(ii) $\phi(x)$ *and* $\phi(y)$ *agree in their first n terms.*

Proof We observe that

$$\phi(I_0) = V_0,\ \phi(I_1) = V_1;$$

$$\phi(I_{00}) = V_{00},\ \phi(I_{01}) = V_{01},\ \phi(I_{10}) = V_{11},\ \phi(I_{11}) = V_{10}.$$

For every word w there is a word $\widetilde{w}$ of same length such that $\phi(I_w) \subset V_{\widetilde{w}}$ and $\phi^{-1}(V_{\widetilde{w}}) = I_w \cap D$. Also $w \longmapsto \widetilde{w}$ is a bijection from $\{0, 1\}^*$ to itself.

From this single observation, three consequences follow immediately:

(1) ϕ is continuous. This is because the pre-image of every basic open set V_w is open in D.
(2) If $\phi(x)$ and $\phi(y)$ belong to V_w and if $|w| = n$, then x and y belong to the same sub-interval I_u for some word u of length n and therefore $|x - y| < \frac{1}{2^n}$. The converse is true as well. In other words the above lemma is proved.
(3) If $\phi(x) = \phi(y)$, then $|x - y| < \frac{1}{2^n}$ holds for all n, and so $x = y$.

This proves that ϕ is one-one. All these together prove that ϕ is a homeomorphism from D_1 to $\phi(D_1) = D_2$.

We next verify that for all $x \in D_1$,

$$\begin{aligned}(\phi(f(x)))_n = 0 \Longleftrightarrow f^n(f(x)) < \frac{1}{2} \Longleftrightarrow f^{n+1}(x) < \frac{1}{2}\\ \Longleftrightarrow (\phi(x))_{n+1} = 0 \Longleftrightarrow (\sigma\phi(x))_n = 0.\end{aligned}$$

This proves : $\phi \circ f = \sigma \circ \phi$. Thus ϕ is a topological conjugacy between the (tent map)$|_{D_1}$ and the (shift map)$|_{D_2}$. Because the latter is known to be topologically transitive, the former has to be. It then follows that the tent map on I itself is topologically transitive. □

Comments on the various proofs: Each proof has its own merit. But they serve the purpose of leading to very general theorems and concepts.

4 Three Ingredients of Chaos

According to a definition proposed by R. Devaney [7], there are three ingredients of chaos:

(T) Topological transitivity (studied in the previous section);
(DP) Dense set of Periodic points;
(SDIC) Sensitive Dependence on Initial Conditions as defined below.

Definition 4.1 *(SDIC)* We say that a dynamical system (X, f) where X is a metric space, has SDIC if there is a positive $\delta > 0$ (called the sensitivity constant) such that $\forall x \in X$ and $\forall r > 0$, $\exists y \in X$, $\exists n \in \mathbb{N}$ such that

$$d(x, y) < r \quad \text{but} \quad d(f^n(x), f^n(y)) > \delta.$$

This means: the orbit of x deviates from that of y by at least δ. Note that we want the same δ at all the points and that the time n of deviation is required only once and that n may vary with x.

In this section, we wish to assert that the three properties T, DP and SDIC are so highly independent, in the sense that

(i) none of them implies any other; and in fact,
(ii) no two of them imply the third.

This means that there is no redundancy in the definition of Devaney's chaos. For this purpose, we actually construct eight examples, one for each row of the following chart:

	T	DP	SDIC
Example 1	✓	✓	✓
Example 2	✓	✓	x
Example 3	✓	x	✓
Example 4	✓	x	x
Example 5	x	✓	✓
Example 6	x	✓	x
Example 7	x	x	✓
Example 8	x	x	x

4.1 T, DP and SDIC

The tent map $f(x) = 1 - |1 - 2x|$ studied in the previous section possesses all these three properties. We in fact proved that for each dyadic sub-interval of length $\frac{1}{2^n}$, it is true that f^n maps that sub-interval onto the whole I. Hence f is topologically transitive. For the same reason, it has SDIC. If $x \in I$ and if $r > 0$ are given, choose a dyadic sub-interval J containing x whose length $\frac{1}{2^n}$ is $< r$. Then because $f^n(J) = I$, there are elements $y, z \in J$ such that $f^n(y) = 1$ and $f^n(z) = 0$. It follows that the orbit of x deviates from that of y or z by at least $\frac{1}{2}$ at this time n. We use triangle inequality here. It follows that any positive $\delta < \frac{1}{2}$ serves as the sensitivity constant. Lastly, we have already seen that all dyadic rational numbers are f-periodic; we know that they form a dense set. Thus the tent map possesses all the three properties T, DP and SDIC.

4.2 T, DP & NOT SDIC

Let $n \in \mathbb{N}$. Let X be the finite set $\{1, 2, \ldots, n\}$ with discrete metric and discrete topology. Let f be the map from $X \longrightarrow X$ defined by $f(x) = x + 1 \pmod{n}$ for all $x \in X$. This f is called an n-cycle. It has only one orbit. The orbit of 1 is $\{1, 2, \ldots, n\}$, the whole set X. Therefore it is topologically transitive. (Highly so, because against the requirement of a dense orbit, we have the whole set as an orbit). All its points are periodic, because $f^n(x) = x$ for all $x \in X$. Therefore f has DP. (Highly so, because against the requirement of a dense set of periodic points, we have the whole set of periodic points). There is no SDIC in this example. It can be seen in two ways:

Proposition 4.2 *At an isolated point, sensitive dependence cannot be there.*

Proof This is because in the ball that is a singleton, we cannot find another point whose orbit deviates from that of the centre ; we cannot find another point at all. □

Proposition 4.3 *An isometry cannot have SDIC.*

Proof An isometry is a map that preserves distances. If r is the radius of a ball around x, then for every y in that ball $d(x, y) < r$ and so for every $n \in \mathbb{N}$, $d(f^n(x), f^n(y)) < r$ (because f^n also preserves distances). If this deviation should be $> \delta$, then that δ should be $< r$. This is true for every $r > 0$. Therefore there is no sensitivity constant. □

Thus we have proved that the n-cycles possesses T and DP but not SDIC. Since this is true for every positive integer n, we have actually infinitely many examples that possess T and DP but not SDIC (but not in I, see Theorem 5.1). But in a later subsection, we shall prove that we do not have an infinite example of this kind. (Note: Infinite example is not same as infinitely many examples).

4.3 T, NOT DP & SDIC

Let f be the restriction of the tent map to the set $\mathbb{Q}^c$ of irrational numbers. Note that $\mathbb{Q}^c$ is f-invariant (If $2x$ or $2 - 2x$ is rational, so should x be). Then f is topologically transitive, by Lemma 3.7.

We next note that we have deliberately omitted all the periodic points of the tent map from the domain of f, so that this example does not have any periodic point at all. Far from DP. That this example does have SDIC follows from the more general result, in the next proposition.

Remark 4.4 Thick subsets are those with non-empty interior.

Proposition 4.5 *Let $g : X \longrightarrow X$ be strongly transitive. (i.e., if V is any thick subset of X, there is $n \in \mathbb{N}$ such that $g^n(V) = X$). Then the restriction of g to any dense invariant subset of X has SDIC. (This means : Not only (X, g) but also every dense subsystem of it has SDIC).*

Proof Any positive δ less than half the diameter of X serves as a sensitivity constant. To prove this let $p, q \in X$ be any two elements. Let Y be a dense subset of X such that $g(Y) \subset Y$. Let $y \in Y$ and $r > 0$. Then the ball $B = B(y, r)$ in X admits $n \in \mathbb{N}$ such that $g^n(B) = X$. It follows that $g^n(B \cap Y)$ is dense in Y (we use the fact that the continuous map g^n takes dense subsets of B to dense subsets of its image). Therefore there are two elements $y, z \in B \cap Y$ such that $g^n(y)$ is near p and $g^n(z)$ is near q, so that the distance between them is very near $d(p, q)$ and hence arbitrarily near the diameter of X. It follows from triangle inequality that the orbit of x deviates from that of y or z by at least half of that $d(p, q)$. □

Thus this example has T and SDIC but not DP.

4.4 T, NOT DP & NOT SDIC

Now we are in search of a transitive map that has neither DP nor SDIC. We cannot find such an example among interval maps. (Later in this chapter we shall prove that T $\Longrightarrow$ DP for interval maps). When we allow other domains, we can find better examples. On the circle $\mathbb{S}^1$, there are topologically transitive maps having neither DP nor SDIC.

Indeed the irrational rotations satisfy our requirements. Let θ be an irrational multiple of π. The rotation by an angle θ is same as the map $f(z) = ze^{i\theta}$ (we use the usual multiplication of complex numbers). That these maps are topologically transitive, follows from a classical theorem of Jacobi (we do not include its proof here, because we confine our interest to real dynamics; but for a proof see subsection (1.2) in [6]).

We can easily prove that there are no periodic points. Indeed, for every positive integer n, $f^n(z) = ze^{in\theta}$. If it were $= z$, then $e^{in\theta} = 1$; this would imply that $n\theta$ is an integral multiple of 2π; and therefore θ is a rational multiple of π; this is contrary to our assumption.

Lastly, we prove that this does not have SDIC. We resort to Proposition 4.3 of this section. We observe that these rotations are isometries. In fact, if z_1 and z_2 are in $\mathbb{S}^1$, $|z_1e^{in\theta} - z_1e^{in\theta}| = |z_1 - z_2||e^{in\theta}| = |z_1 - z_2|$.

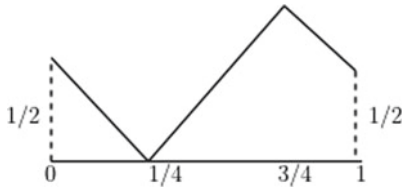
Thus this example provides infinitely many examples of systems that are topologically transitive, without DP and without SDIC.

4.5 NOT T, DP & SDIC

Let f be the piecewise linear map on I specified by

$$f(0) = \frac{1}{2}; \quad f\left(\frac{1}{4}\right) = 0; \quad f\left(\frac{3}{4}\right) = 1; \quad f(1) = \frac{1}{2}.$$

Its graph is as sketched here.



Its formula is

$$f(x) \;=\; \begin{cases} \frac{1}{2} - 2x & \text{if } 0 \le x \le \frac{1}{4}; \\ 2x - \frac{1}{2} & \text{if } \frac{1}{4} < x \le \frac{3}{4}; \\ \frac{5}{2} - 2x & \text{if } \frac{3}{4} < x \le 1. \end{cases}$$

This map is not topologically transitive because $[0, \frac{1}{2}]$ is f-invariant. We now claim that every number in I that is of the form $\frac{2m+1}{4n+2}$ (where $m, n \in \mathbb{N}$) is f-periodic. Note that these are the numbers

$$\frac{1}{2}, \frac{1}{6}, \frac{5}{6}, \frac{1}{10}, \frac{3}{10}, \frac{7}{10}, \frac{9}{10}, \frac{1}{14}, \ldots.$$

In this sequence of numbers, we can verify that the first three are fixed points, the next four are points of period 2, and so on.

But, for a rigorous proof, we have to argue more cleverly. For a fixed positive integer n, let

$$A_n \;=\; \left\{ \frac{2m+1}{4n+2} \;:\; 0 \le m \le 2n \right\}.$$

This is a finite set having $2n + 1$ elements. We easily see that this set is f-invariant, because if $x = \frac{2m+1}{4n+2}$, then the numbers $\frac{1}{2} - 2x$, $2x - \frac{1}{2}$ and $\frac{5}{2} - 2x$ are all of this form (as can be verified by direct calculation). On the other hand every such number (i.e., $x = \frac{2m+1}{4n+2}$) is the image (under f) of another such number namely

$$\begin{aligned} y &= \frac{m-n}{4n+2} && \text{if } m < 2n \text{ and if } m \text{ and } n \text{ are different parity;} \\ \text{or } \; y &= \frac{m+n+1}{4n+2} && \text{if } m \text{ and } n \text{ are same parity;} \\ \text{or } \; y &= \frac{5n-m+2}{4n+2} && \text{if } m \ge 2n \text{ and if } m \text{ and } n \text{ are different parity.} \end{aligned}$$

Thus f is a bijection on each A_n. We use now:

Proposition 4.6 *Let (X, g) be a dynamical system. Let $F \subset X$ be a finite subset such that $g(A) = A$. Then every element of A is g-periodic.*

Proof $g|_A$ becomes a permutation and hence has a cyclic decomposition. This exhibits the cyclic orbits of all elements of A. □

We now continue with our present example. Every element of A_n is f-periodic, for every $n \in \mathbb{N}$. Next we note that A_n is a $\frac{1}{n}$-net in I (in the sense that every element in I is at a distance $< \frac{1}{n}$ from some element in A). This is because in $A_n = \left\{\frac{1}{4n+2}, \frac{3}{4n+2}, \ldots, \frac{4n+1}{4n+2}\right\}$ any two adjacent elements are at a distance $\frac{1}{2n+1}$, and they are equally spread.

Thus the set of f-periodic points is a countable union of finite sets that are $\frac{1}{n}$-nets, for each positive integer n. Therefore it follows immediately that this set is dense in I.

Remark 4.7 1. There are other proofs for the same result. But this proof has an advantage of being elementary and self-contained.
2. One can prove that there are no other f-periodic points in this example. But we do not need this result here.

Lastly, we now prove that this has SDIC. For this, we take a different approach. We look at the backward orbit of the three fixed points $\frac{1}{2}$, $\frac{1}{6}$ and $\frac{5}{6}$. Here, backward orbit of p means the set of all points that contain p in the orbit. We claim:

(a) the backward orbit of $\frac{1}{2}$ is dense in $[0, 1]$.
(b) the backward orbit of $\frac{1}{6}$ is dense in $[0, \frac{1}{2}]$.
(c) the backward orbit of $\frac{5}{6}$ is dense in $[\frac{1}{2}, 1]$.

Proof of (a): We find that $0, 1$ go to $\frac{1}{2}$ at time 1; $\frac{1}{4}$ and $\frac{3}{4}$ go to $\frac{1}{2}$ at time 2; four more points $\frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}$ go to $\frac{1}{2}$ at time 3 and so on. In other words, the sets $f^{-1}\left(\frac{1}{2}\right), f^{-2}\left(\frac{1}{2}\right), f^{-3}\left(\frac{1}{2}\right)$ etc. are all finite sets and their cardinality is increasing.

We next note that $f^{-1}\left(\frac{1}{2}\right)$ is a $\frac{1}{4}$-net; $f^{-2}\left(\frac{1}{2}\right)$ is a $\frac{1}{8}$-net and so on. So, their union (which is the same as the backward orbit of $\frac{1}{2}$) is an ϵ-net for all $\epsilon > 0$. So, it is a dense set.

Proof of (b): We find that

$$f^{-1}\left(\frac{1}{6}\right) = \left\{\frac{1}{3}, \frac{1}{6}\right\};\ f^{-2}\left(\frac{1}{6}\right) \supset \left\{\frac{1}{12}, \frac{5}{12}\right\};$$

$$f^{-3}\left(\frac{1}{6}\right) \supset \left\{\frac{1}{24}, \frac{5}{24}, \frac{7}{24}, \frac{11}{24}\right\};\ \text{and so on.}$$

Thus, the backward orbit of $\frac{1}{6}$ includes all points of the form $\frac{2m+1}{3 \cdot 2^n}$ (after eliminating common factors) that are $< \frac{1}{2}$ (where $m, n \in \mathbb{N}_0$). As before, we can prove that this is a dense subset of $[0, \frac{1}{2}]$.

Proof of (c): It is similar to that of (b) and hence omitted.

Now we complete the proof of SDIC as follows.

Proposition 4.8 *Let (X, f) be a dynamical system where every thick set V has two points whose orbits deviate by at least $\delta > 0$. Then $\frac{\delta}{2}$ serves as a sensitivity constant for (X, f).*

Proof Let $x \in X$ and $r > 0$. By assumption the ball $B(x, r)$ has two points y, z such that $d(f^n(y), f^n(z)) > \delta$ for some $n \in \mathbb{N}$. Now by triangle inequality, either $d(f^n(x), f^n(y))$ or $d(f^n(x), f^n(z))$ is bigger than $\frac{\delta}{2}$. Thus (X, f) has SDIC with $\frac{\delta}{2}$ as sensitivity constant. □

We now return to our example. If V is any thick subset of I, V should meet the backward orbit of $\frac{1}{2}$ (because of (a)); it should also meet the backward orbit of $\frac{1}{6}$ or that of $\frac{5}{6}$ (because of (b) and (c)). Thus there are points $x, y \in V$ and $m, n \in \mathbb{N}$ such that

$$f^m(x) = \frac{1}{2} \quad \text{and} \quad f^n(y) \in \left\{\frac{1}{6}, \frac{5}{6}\right\}.$$

Because $\frac{1}{2}$, $\frac{1}{6}$ and $\frac{5}{6}$ are fixed points, if $k >$ both m and n, then $f^k(x) = \frac{1}{2}$ and $f^k(y) \in \left\{\frac{1}{6}, \frac{5}{6}\right\}$. So the distance between them is

$$|f^k(x) - f^k(y)| = \left(\left|\frac{1}{2} - \frac{1}{6}\right| \text{ or } \left|\frac{1}{2} - \frac{5}{6}\right|\right) = \frac{1}{3}.$$

Thus, the orbits of x and y deviate by a distance of $\frac{1}{3}$. So by Proposition 4.8, (I, f) has SDIC where any positive number $< \frac{1}{6}$ works as a sensitivity constant. Thus we have proved that this example has DP and SDIC but it is not transitive.

Remark 4.9 Another method to prove this is to prove a stronger result stated below.

Proposition 4.10 *If f is as the example described above, then $f|_{[0,\frac{1}{2}]}$ is topologically conjugate to the tent map, via the homeomorphism $\frac{1-x}{2}$; similarly $f|_{[\frac{1}{2},1]}$ is topologically conjugate to the tent map via the homeomorphism $\frac{1+x}{2}$.*

We omit the proof because this result is dispensable now.

4.6 NOT T, DP & NOT SDIC

The identity map on I has DP (because all points are fixed points) but is not topologically transitive (because every subset is invariant) and does not have SDIC (because it is an isometry). The reflection map $1 - x$ has DP (because every point is a periodic point) but is not topologically transitive (because $\left[\frac{1}{4}, \frac{3}{4}\right]$ is invariant) and does not have SDIC (because it is an isometry).

4.7 NOT T, NOT DP & SDIC

Consider the system of example described in subsection NOT T, DP & SDIC. We need two results about it (for discussing this seventh example).

(i) Every periodic point for that f, is a rational number.
(ii) Both $\mathbb{Q}$ and $\mathbb{Q}^c$ are f-invariant.

For proving (i), we first observe that every compositional power f^n is a piecewise linear map, where in every piece, it has a formula of the form $f^n(x) = ax + b$ where a and b are rational numbers and such that $|a| = 2^n$. Therefore if x is a periodic point of f, then $x = f^n(x)$ for some $n \in \mathbb{N}$ and therefore $x = ax + b$ for some rational $a,\ b$. This implies that $x = \frac{b}{1-a}$ is itself a rational number.

To prove (ii), we observe two things :

(a) If x is rational, then the numbers $\frac{1}{2} - 2x,\ 2x - \frac{1}{2}$ and $\frac{5}{2} - 2x$ are rational.
(b) If $\frac{1}{2} - 2x$ or $2x - \frac{1}{2}$ or $\frac{5}{2} - 2x$ is rational, then x is rational.

Now we are ready to describe our desired example. It is nothing but the restriction of f (in the example described in subsection NOT T, DP & SDIC) to the set of all irrational numbers (whose invariance was noted just now); call it g. We have just now proved in (i) that g has no periodic points at all. g is not topologically transitive because the set $[0, \frac{1}{2}] \cap \mathbb{Q}^c$ is invariant. But, g has SDIC because of the next result.

Proposition 4.11 *Let (X, f) have SDIC and let Y be a dense f-invariant subset of X. Then $(Y, f|_Y)$ also has SDIC.*

Proof Let $\delta > 0$ be a sensitivity constant of (X, f). Let $y \in Y$ and $r > 0$. Then $\exists z \in B(y, r)$ and $n \in \mathbb{N}$ such that $d(f^n(y), f^n(z)) > \delta$. But this z may not be in Y (and we are looking for one such element in Y). Because f^n is a continuous map, the function $z \longmapsto d(f^n(y), f^n(z))$ is also continuous from X to $\mathbb{R}$. Therefore there is $s > 0$ such that for all w in $B(z, s)$ the number $d(f^n(y), f^n(w)) > \delta$. One such w can be chosen in Y because the dense set Y should meet $B(z, s)$. We have proved that the same δ serves as a sensitivity constant for (Y, f) also. □

Remark 4.12 The converse of the Proposition 4.11 is also true. (X, f) has SDIC if and only if $(Y, f|_Y)$ has SDIC. The condition that Y is dense in X cannot be omitted.

Now we return to our example in this section. We have completed the proof of the fact it has SDIC but is neither T nor DP for this example.

4.8 NOT T, NOT DP & NOT SDIC

Let $f(x) = x^2$ on $[0, 1]$. Then f is not topologically transitive because $[0, \frac{1}{2}]$ is f-invariant. Also f does not have DP because 0 and 1 are the only periodic points. To prove this we use the following result:

Proposition 4.13 *For an increasing interval map, all periodic points are fixed points.*

Proof If p is not a fixed point, then $p < f(p)$ or $p > f(p)$. Without loss of generality assume $p < f(p)$. Because f is increasing, this implies $f(p) < f^2(p)$. Recursively we have $p < f(p) < f^2(p) < \cdots$. It is not possible to have $p = f^n(p)$ for any $n \in \mathbb{N}$. So, p is not a periodic point. □

We now return to our example. It does not have SDIC because of the next result.

Proposition 4.14 *In a dynamical system (X, f) where X is a metric space, we have*

(a) SDIC is inherited by open invariant sets.
(b) No contraction map has SDIC.

Proof **Proof of (a):** Let (X, f) have SDIC and let Y be an open f-invariant subset of Y. To prove that $(Y, f|_Y)$ is also having SDIC, let $y \in Y$ and $r > 0$. We may assume that the ball $B = B(y, r) \subset Y$. Because (X, f) has SDIC, $\exists z \in B$ and $n \in \mathbb{N}$ such that $d(f^n(z), f^n(y)) > \delta$, the sensitivity constant. This proves that $(Y, f|_Y)$ also has SDIC.

Proof of (b): Let (X, f) be a contraction map. If $x, y \in X$, then $d(f^n(x), f^n(y)) \leq c^n d(x, y)$ where c is the contraction constant. Now let $\delta > 0$ be,if possible, a sensitivity constant for (X, f). First, no point can be isolated, because in its singleton-neighborhood, sensitivity would fail. Take any element $x \in X$ and any other element y in $B(x, \delta)$. Then the trajectory of y can not deviate from that of x by more than δ because it is a contraction. □

We again return to our example. We claim that it is a contraction map on the invariant open set $[0, \frac{1}{3})$. In fact if $0 \leq x, y < \frac{1}{3}$, then $|x^2 - y^2| = |x - y||x + y| \leq \frac{2}{3}|x - y|$. Thus $\frac{2}{3}$ serves as a contraction constant. Since $[0, \frac{1}{3})$ is open and invariant in $[0, 1]$, it follows that x^2 on $[0, 1]$ does not have SDIC. Thus, we have completed the proof that x^2 has none of the three properties T, DP and SDIC.

Summary of this section : We have seen eight interesting examples of dynamical systems to understand in eight ways that the three properties T, DP and SDIC have no implications among them. In the course of describing these examples and their dynamical properties, we came across ten elementary propositions that are interesting in their own right. But now a question comes up. Among these eight examples, only four are interval maps. Why can't we give all the examples as interval maps? The fact is that the other four cannot be interval maps. In other words, among interval maps, these three properties are not independent, because T implies the other two. (This we shall prove soon). Whenever possible, we have provided interval maps as examples. Had it not been so, there are easier examples in place of the example described in subsection NOT T, DP & SDIC. For instance, the **union** of the two tent maps on two disjoint intervals.

5 Chaos For Interval Maps

5.1 For Interval Maps Transitivity Implies Chaos

Contrary to what we saw in the previous section, now we are going to prove that, in some classes of dynamical systems, there are striking implications among the three properties T, DP and SDIC.

Theorem 5.1 *$T \Longrightarrow$ DP is true for all interval maps.*

Theorem 5.2 ([4]) *$T \wedge DP \Longrightarrow$ SDIC is true on all infinite metric spaces; in other words every infinite metric space possessing T and DP should satisfy SDIC.*

Before proving these, we prove some results on trajectory behaviour that are interesting independently.

Theorem 5.3 *Let $m < n$ and let f be an interval map and let x be a point such that $x < f^n(x) < f^m(x)$. Then there is some f-periodic point between x and $f^m(x)$.*

Proof Consider the f^m-trajectory of $f^m(x)$. Ask whether their terms are $< f^m(x)$ or not. If $f^{2m}(x) < f^m(x)$, then under f^m, x has moved to the right side, and $f^m(x)$ has moved to the left side, and so there should be a fixed point of f^m between them; this point is a f-periodic point.

In the other case $f^{2m}(x) \geq f^m(x)$. Suppose

$$f^{km}(x) \;<\; f^m(x) \quad (\text{with } f^{(k-1)m}(x) \;\geq\; f^m(x)),$$

then under the map $f^{(k-1)m}$, x moves to the right side and $f^m(x)$ moves to the left side, and so, there should be a fixed point of $f^{(k-1)m}$ between them; this again gives a f-periodic point.

Thus we are able to prove the result, except when the entire trajectory of $f^m(x)$ under f^m is on the right side of $f^m(x)$.

Similarly letting $r = n - m$, considering $f^m(x)$ as the initial point, and looking at its trajectory under f^r, we can prove that there is a f-periodic point between $f^n(x)$ and $f^m(x)$ except when this entire trajectory lies on the left side of $f^n(x)$. Thus, the only remaining case is when $f^{km}(x) \geq f^m(x)$ and $f^{kr}(f^m(x)) \leq f^n(x)$ for all k, where $r = n - m$. But this case cannot arise because $f^{m(r+1)}(x)$ then would be both $\geq f^m(x)$ and $\leq f^n(x)$ which is not possible. □

The result in the above theorem seems deceptively simple, but it has far-reaching consequences. We list four of them below.

Corollary 5.4 *Let J be an open interval without any f-periodic point. Then for elements $x \in J$, all orbit-portions in J are monotonic.*

Proof Admittedly, some terms of the orbit of x can go outside J; admittedly, since no assumption on f has been made outside J, the orbit-behaviour can be arbitrary there; admittedly, at some future time it can come back to J. But whenever they come back to J they have to follow a discipline as stated below. If $n_1 < n_2 < \cdots$ are the time instants when x visits J, we have that the sequence $x, f^{n_1}(x), f^{n_2}(x), \ldots$ should be monotone. There are two possibilities:

1. Either $x < f^{n_1}(x) < f^{n_2}(x) < \cdots$;
2. or $x > f^{n_1}(x) > f^{n_2}(x) > \cdots$.

This portion of the orbit of x in J may be finite or infinite. □

If we examine the above proof carefully, we can even estimate the period of the f-periodic point that is guaranteed in Theorem 5.3. Because the two arithmetic progressions $m, 2m, 3m, \ldots$ and $n, 2n - m, 3n - 2m, \ldots$ have $m + \text{lcm}(m, n - m)$ as a common term, we are sure that some f-periodic point between x and $f^m(x)$ has f-period that divides $m + \text{lcm}(m, n - m)$. Here are some particular instances.

Corollary 5.5 • *If $x < f^2(x) < f(x)$, then there is a fixed point between x and $f(x)$.*
- *If $x < f^3(x) < f(x)$, then there is a point of period 1 or 2 between x and $f(x)$.*
- *If $x < f^3(x) < f^2(x)$, then there is a periodic point $f^4(y) = y$ between x and $f^2(x)$.*
- *If $x < f^4(x) < f^2(x)$, then there is a periodic point $y = f^4(y)$ between x and $f^2(x)$.*
- *If $x < f^4(x) < f^3(x)$, then there is a periodic point $y = f^6(y)$ between x and $f^3(x)$.*
- *If $x < f^5(x) < f^2(x)$, then there is a periodic point $y = f^{12}(y)$ between x and $f^2(x)$.*

Corollary 5.6 *If an open interval J contains no f-periodic point, then J meets any omega-limit set at at-most one point.*

Proof Let $x \in I$ be such that some $y \in J \cap \omega(f; x)$. Then infinetely many terms of $(f^n(x))$ are in J. They form a monotone sequence and cannot converge to two distinct elements. □

Corollary 5.7 *If J is an open interval containing no f-periodic points, then J contains no f-recurrent points.*

Proof If x is f-recurrent, then the trajectory of x has a subsequence that converges to x. But by Theorem 5.3, the trajectory of x has its portion in J monotonic, moving farther and farther away from x. □

Corollary 5.8 $\overline{P(f)} = \overline{R(f)}$ *for every interval map. [Here $P(f)$ denotes the set of all f-periodic points; $R(f)$ denote the set of all f-recurrent points].*

Proof Obviously $\overline{P(f)} \subset \overline{R(f)}$. To prove the reverse inequality, it suffices to prove that every open interval J disjoint from $P(f)$ is disjoint from $R(f)$ also. But this is what is stated in Corollary 5.7. □

Corollary 5.9 *$T \Longrightarrow$ DP for interval maps. (This is Theorem 5.1 of this section)*

Proof If an interval map f is transitive, it has a dense orbit. All elements in this dense orbit are recurrent. Therefore $\overline{R(f)} = I$. Therefore by Corollary 5.8, $\overline{P(f)} = I$. This means f has DP. □

5.2 *T & DP $\Longrightarrow$ SDIC*

Theorem 5.10 ([4, 16]) *Let X be an infinite metric space. Let $f : X \longrightarrow X$ be a topologically transitive map with a dense set of periodic points. Then f has SDIC.*

Proof Let $\delta > 0$ be the distance between two periodic f-orbits. (Because f has DP and because X is infinite, there should be infinitely many periodic orbits; any two distinct periodic orbits are disjoint; choose any two of them; there is positive distance between their orbits, as it is so, for any two disjoint finite sets).

For every x in X, one of these two periodic orbits has to be at a distance $> \frac{\delta}{2}$ from x. (If both are at a distance $< \frac{\delta}{2}$ from x, then the triangle inequality will make the distance between these orbits as $< \delta$). We shall prove that $\frac{\delta}{8}$ serves as a sensitivity constant for f.

Let $r > 0$. We may assume $r < \frac{\delta}{8}$. Choose a periodic point p in $B(x, r)$ and the cyclic orbit $q, f(q), \ldots$, that is at a distance $> \frac{\delta}{2}$ from x. Let k be the f-period of p. Let

$$W = \left\{ y \in X \mid d(f^i(y), f^i(q)) < \frac{\delta}{8} \text{ for } 0 \leq i \leq k \right\}.$$

Because each f^i is continuous, W is an open neighborhood of q. Because f is topologically transitive there is some z in $B(x, r)$ and some $n \in \mathbb{N}$ such that $f^n(z) \in W$. This implies $f^{n+i}(z)$ is at a distance $< \frac{\delta}{8}$ from the f-orbit of q, for $0 \leq i \leq k$. One of these $n + i$'s has to be a multiple of k (as there are k consecutive terms). For that i, we have $f^{n+i}(p) = p$ and $f^{n+i}(z)$ is at a distance $< \frac{\delta}{8}$ from the orbit of q. Therefore,

$$\begin{aligned} \frac{\delta}{2} &< d(x, f^{n+i}(q)) \leq d(x, p) + d(p, f^{n+i}(z)) + d(f^{n+i}(z), f^{n+i}(q)) \\ &< \frac{\delta}{8} + d(p, f^{n+i}(z)) + \frac{\delta}{8}. \end{aligned}$$

This gives $d(p, f^{n+i}(z)) > \frac{\delta}{4}$. Thus the orbits of p and z deviate from each other at least by $\frac{\delta}{4}$. It follows that the orbit of x should deviate from one of them (either

orbit p or obit of z) by at least half of it, namely $\frac{\delta}{8}$. This proves that $\frac{\delta}{8}$ becomes the sensitivity constant. □

6 Some Consequences of Intermediate Value Theorem in Dynamics

We use the abbreviation IVT for Intermediate Value Theorem. This theorem states:

Theorem 6.1 *If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, and if s is any number between $f(a)$ and $f(b)$, then there is some c between a and b such that $f(c) = s$.*

6.1 Immediate Applications

First we state the main results of this section, each of which can be proved by clever (and sometimes repeated) applications of IVT. Proofs are left as exercises.

Theorem 6.2 *If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, and if f moves some point p to its left, and some point q to its right, then f has a fixed point between p and q.*

Theorem 6.3 (Fixed Point Theorem) *Every interval map has a fixed point.*

Theorem 6.4 *If $f : [a, b] \longrightarrow [c, d]$ is continuous, surjective and if $[a, b] \subset [c, d]$, then f has a fixed point.*

Theorem 6.5 *If I and J are two closed intervals and f is a continuous real map such that $f(I) \supset J$, then there is a closed sub-interval $K \subset I$ such that $f(K) = J$.*

Theorem 6.6 *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous. Let $p, q, c, d \in \mathbb{R}$ be such that $p < q$ are in the domain of f and let $c < d$. Then the following are equivalent.*

(1) *$[p, q]$ is a minimal interval satisfying $f([p, q]) = [c, d]$.*
(2) *$f^{-1}(\{c, d\}) \cap [p, q] = \{p, q\}$.*

6.2 Sarkovskii's Theorem: A Statement

Notation 6.7 Let $m, n \in \mathbb{N}$. We write $m \succ n$ if m precedes n in the following total order:

$$
\begin{aligned}
& 3 \succ 5 \succ 7 \succ \cdots \\
\succ\ & 3 \cdot 2 \succ 5 \cdot 2 \succ 7 \cdot 2 \succ \cdots \\
\succ\ & 3 \cdot 2^2 \succ 5 \cdot 2^2 \succ 7 \cdot 2^2 \succ \cdots \\
& \vdots \\
\succ\ & \cdots \succ 2^3 \succ 2^2 \succ 2 \succ 1.
\end{aligned}
$$

This is called the Sarkovskii's ordering of $\mathbb{N}$. We use the same symbol for the reflexive, transitive relation generated by it. (i.e., $3 \succ 3$ and $3 \succ 7$ etc. are true.)

Definition 6.8 Let $m, n \in \mathbb{N}$. We say that m forces n if every continuous $f : \mathbb{R} \longrightarrow \mathbb{R}$ that admits a periodic point of period m has to (necessarily) admit a periodic point of period n.

Theorem 6.9 (Sarkovskii) *Let $m, n \in \mathbb{N}$. Then m forces n if and only if $m \succ n$.*

Corollary 6.10 3 *forces every other positive integer.*

Theorem 6.11 *Let A be a nonempty subset of $\mathbb{N}$. Then $A = per(f)$ for some continuous $f : \mathbb{R} \longrightarrow \mathbb{R}$ if and only if A satisfies:*

$$m \in A \text{ and } m \succ n \implies n \in A.$$

Here, $per(f)$ denotes the set of all periods of f-periodic points.

Example 6.12 • $\{1, 2, 2^2, \ldots\} = \{\text{powers of } 2\}$ is one such set.
• $\{1\} \cup \{\text{even positive itegers}\}$ is another such set.
• $\{\text{odd integers}\}$ is not such a set.

Corollary 6.13 *(1) If $per(f)$ is finite, then every element in it has to be a power of 2.*
(2) If $per(f)$ contains an odd integer, then its complement in $\mathbb{N}$ is finite.

Remark 6.14 When we characterise the sets of periods for real maps (after knowing the sets of periods of interval maps) the only difference is that the empty set can be $per(f)$. In other words,

$$\{per(f) \mid f \text{ is a real map }\} \ = \ \{\varnothing\} \cup \{per(f) \mid f \text{ is an interval map }\}.$$

6.3 Digraphs of Cycles

Recall that we use the notation $I \xrightarrow{f} J$ to mean that I and J are closed intervals such that $f(I) \supset J$. Sometimes, when the function f is clear from the context we simply write $I \longrightarrow J$.

A cycle in a graph G is a directed path whose starting point is same as the ending point. For instance in the graph of Example 6.17,

1. $I_3 \longrightarrow I_3$ is a cycle of length 1.
2. $I_2 \longrightarrow I_3 \longrightarrow I_2$ is a cycle of length 2.
3. $I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_1$ is a cycle of length 3.
4. $I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_3 \longrightarrow I_1$ is a cycle of length 4.
5. $I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_1$ is a cycle of length 5, and so on.

A word of caution: We are using the same word 'cycle' in two different senses (one in graph theory and other in dynamics); unfortunately, both these senses are coming together in our present discussion; we have to carefully interpret them. For instance, in Example 6.17, we started with an f-orbit that was a 4-cycle. To this 4-cycle, we associated a digraph. In that digraph we saw examples of cycles of different lengths. These are graph-cycles in a graph associated to a dynamical cycle.

Another word of caution: The arrow symbol is also used in two different senses. When we write $f : [a, b] \longrightarrow [c, d]$ we mean that f is a function whose domain is $[a, b]$ and whose codomain is $[c, d]$. Here, $f([a, b])$ need not contain $[c, d]$. (On the contrary, $f([a, b]) \subset [c, d]$). But when we write $[a, b] \xrightarrow{f} [c, d]$ (in this section, while describing the edges of the graph) we mean that $f([a, b]) \supset [c, d]$.

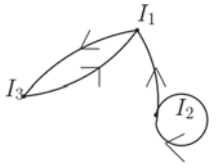
We do not want to avoid this notational confusion, because these notations have become standard and because they are convenient; a closer look will avoid the confusion.

Let f be a continuous map from $[a, b]$ to itself. We associate a digraph to every f-cycle as follows.

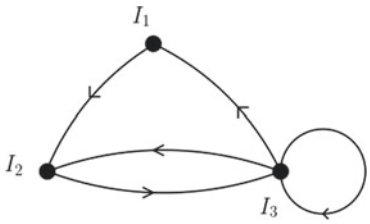
Let $x_1 < x_2 < \cdots < x_n$ be the elements in a cyclic orbit of f. We denote $I_j = [x_j, x_{j+1}]$ for $1 \leq j \leq n-1$. Each of these intervals is taken as a vertex of a graph G. Thus G has $n-1$ vertices. We draw an edge from a vertex I_j to a vertex I_k if $f(I_j) \supset I_k$. In this manner, we obtain a directed graph.

Example 6.15 If f has a 3-cycle $a = f^3(a) < f(a) < f^2(a)$, then the digraph of this 3-cycle has 2 vertices namely $I_1 = [a, f(a)]$ and $I_2 = [f(a), f^2(a)]$. We note that $f(I_1) \supset I_2$ (by IVT) and that $f(I_2) \supset I_1 \cup I_2$. Therefore in the associated directed graph, we have three directed edges (I_1, I_2), (I_2, I_1) and (I_2, I_2).

Example 6.16 If f has a 4-cycle $a < b < c < d$ such that $f(a) = d$, $f(d) = b$, $f(b) = c$ and $f(c) = a$, then its associated digraph has 3 vertices namely, $I_1 = [a, b]$, $I_2 = [b, c]$ and $I_3 = [c, d]$ and there are 4 directed edges.



Example 6.17 If f has a 4-cycle $a < b < c < d$ such that $f(a) = b$; $f(b) = c$; $f(c) = d$; $f(d) = a$, then the directed graph associated with this 4-cycle has 3 vertices and 5 directed edges as shown in this picture.



6.4 Use of digraphs in the proof of Sarkovskii's theorem

Theorem 6.18 *Let under an interval map f,*

$$I_0 \longrightarrow I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow \cdots \longrightarrow I_n \longrightarrow I_0$$

be a cycle (these intervals need not be distinct) of length $n+1$. Then $\exists x \in I_0$ such that $f^k(x) \in I_k$ for all $1 \leq k \leq n$ and such that $f^{n+1}(x) = x$.

Proof Choose a sub-interval J_n of I_n such that $f(J_n) = I_0$. Proceeding backwards, choose a sub-interval J_{n-1} of I_{n-1} such that $f(J_{n-1}) = J_n$. In choosing these, we are using Theorem 6.5. At the end of this backward succession, choose a sub-interval J_0 of I_0 such that $f(J_0) = J_1$. Then we have

$$J_0 \longrightarrow J_1 \longrightarrow J_2 \longrightarrow \cdots \longrightarrow J_n \longrightarrow I_0.$$

This gives $J_0 \xrightarrow{f^{n+1}} I_0$. Because $J_0 \subset I_0$, Theorem 6.4 applies; we have $x \in J_0$ which is a fixed point of f^{n+1}. This x has the property $f(x) \in J_1$. (This is because $f(J_0) = J_1$); and so on. Since each $J_k \subset I_k$, we arrive at the stated conclusion. □

Remark 6.19 The above theorem can be stated as follows. For every cyclic path in the graph, there is a periodic point whose orbit traces that same path.

This simple-looking result has some profound consequences. We can use it to prove Sarkovskii's theorem (Theorem 6.9). We will not include the full proof here, but all the key ideas needed for the proof will be explained now.

Theorem 6.20 *Let $m, n \in \mathbb{N}$. Consider the digraphs associated with m-cycles of an interval map. (Different m-cycles give rise to different digraphs.) Suppose each of them contains simple n-cycle (in graph-theoretic sense, as stated below). Then m forces n.*

Proof This follows from the previous Theorem.

In this context, a simple n-cycle is a graph-cycle of length n such that the whole cycle is not the union of two or more sub-cycles of the same kind. For instance, $I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_1 \longrightarrow I_2 \longrightarrow I_3$ is not simple, whereas $I_1 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_2 \longrightarrow I_3 \longrightarrow I_1$ is simple. Note that we are allowing repetitions of sub-cycles, even within a simple cycle. What motivates this kind of definition of a simple cycle, is the following: If a periodic point traverses the path of a simple cycle, then its period equals to whole length of the cycle, and is not less. □

Remark 6.21 The power of the above theorem can be appreciated through the following two particular instances:

(1) 3 forces every positive integer.
(2) 4 forces 2 and 1 only.

Proof We prove both the statements in the remark now.

Proof of (1): To prove this, we first note that there are only two possible patterns for a 3-cycle namely:

(i) $a < f(a) < f^2(a)$ with $f^3(a) = a$;
(ii) $a > f(a) > f^2(a)$ with $f^3(a) = a$.

If it is of the type $f(a) < a < f^2(a)$ with $f^3(a) = a$, then taking $b = f^2(a)$, we bring it to the form $b > f(b) > f^2(b)$ with $f^3(b) = b$. In this manner all the 3-cycle patterns can be brought to one of the above two forms. We assume the pattern as $a < f(a) < f^2(a)$ with $f^3(a) = a$.

The other case can be dealt with similarly. Take $I_1 = [a, f(a)]$ and $I_2 = [f(a), f^2(a)]$. The digraph has two vertices and three directed edges. Here

1. $I_2 \longrightarrow I_2$ is a cycle of length 1;
2. $I_1 \longrightarrow I_2 \longrightarrow I_1$ is a cycle of length 2;

3. $I_1 \longrightarrow I_2 \longrightarrow I_2 \longrightarrow I_1$ is a cycle of length 3;
4. $I_1 \longrightarrow I_2 \longrightarrow I_2 \longrightarrow I_2 \longrightarrow I_1$ is a cycle of length 4 and so on.

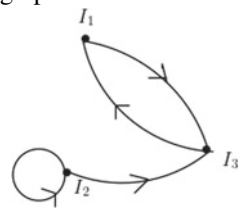
For $n > 2$, by repeating I_2, $(n-1)$ times, we obtain a cycle of length n. The previous theorem now proves that (because in this graph simple n-cycles are available for any $n \in \mathbb{N}$) 3 forces n for all $n \in \mathbb{N}$.

Proof of (2): Among the different patterns of 4-cycles, we have already listed two of them as Examples 6.16 and 6.17; we can prove that every pattern of 4-cycle is order-isomorphic to one of these three:

1. $a < f^2(a) < f(a) < f^3(a)$ with $f^4(a) = a$;
2. $a < f(a) < f^2(a) < f^3(a)$ with $f^4(a) = a$ and
3. $a < f^3(a) < f(a) < f^2(a)$ with $f^4(a) = a$.

For example, if $a < f^3(a) < f^2(a) < f(a)$, then by letting $b = f(a)$, we find the pattern of the orbit of b as exact dual of the third case above.

For the first of them, the digraph is as under:

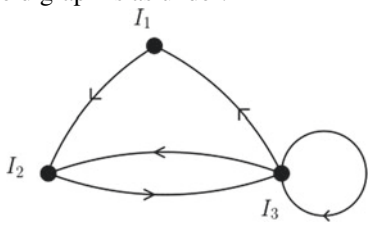


Here,

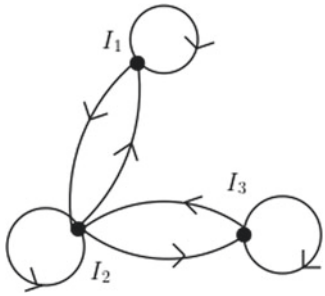
1. $I_2 \longrightarrow I_2$ is a cycle of length 1;
2. $I_1 \longrightarrow I_2 \longrightarrow I_1$ is a cycle of length 2.
3. We find no other simple cycle.

This proves that the number 4 forces no numbers other than 1 and 2.

For the third of them, the digraph is as under:



1. $I_3 \longrightarrow I_3$ is a cycle of length 1;
2. $I_2 \longrightarrow I_3 \longrightarrow I_2$ is a cycle of length 2.
3. There are cycles of greater length as well.



For the fourth of them, the digraph is as under:

1. Here there are 3 cycles of length 1 and 2 cycles of length 2.
2. There are cycles of greater length as well.

Now we summarise our observations: For every 4-cycle, whatever its pattern be, its digraph admits a cycle of length 1 and also a cycle of length 2. This proves that 4 forces both 2 and 1. □

In some of these patterns, we find that the digraph admits longer simple cycles (in fact, of any length). But this yields no further information useful to the present problem.

6.5 *Doubling periods*

For an interval map f, let $per(f)$ denote the set of all periods of periodic points of f. For example, if $f(x) = x^2$, then $per(f) = \{1\}$. If $f(x) = 1 - x$, then $per(f) = \{1, 2\}$. If $f(x) = 1 - |1 - 2x|$, then $per(f) = \mathbb{N}$. For each interval map f, we now associate another interval map $\widetilde{f}$ such that

$$per(\widetilde{f}) \;=\; \{1\} \cup \{2n \mid n \in per(f)\}.$$

Define $g : [0, 3] \longrightarrow [0, 3]$ by

$$g(x) \;=\; \begin{cases} f(x) + 2 & \text{if } 0 \le x \le 1, \\ (2 - x)(2 + f(1)) & \text{if } 1 < x < 2, \\ x - 2 & \text{if } 2 \le x \le 3. \end{cases}$$

Then g is continuous, because the adjacent formula-pieces agree at the common points. g is a linear polynomial on $[1, 2]$ and also on $[2, 3]$. Moreover

$$g([0, 1]) \subset [2, 3]\; g([2, 3]) = [0, 1] \;\text{ and }\; g([1, 2]) = [0, f(1) + 2].$$

Further, $g(g(x)) = f(x) + 2 - 2 = f(x)$ if $0 \le x \le 1$.

Shall we say that f is the square of g on $[0, 1]$ and the set $[0, 1] \cup [2, 3]$ is g-invariant? Because $g([1, 2])$ contains $[1, 2]$, g has a fixed point in that middle interval.

The slope of g in $[1, 2]$ lies between -2 and -3. We use this fact to prove the following: If p is the unique fixed point of g in $[1, 2]$ and if x is any other point there, then $|g(x) - p| \ge 2|x - p|$. Thus $g(x)$ is at least twice farther away from p than x is. If $g(x)$ is also in $[1, 2]$, then the same applies again to yield that $g(g(x))$ is much farther from p. We conclude that the g-orbit of x must leave $[1, 2]$ at some time or other. i.e., $\exists n \in \mathbb{N}$ such that $g^n(x) \notin [1, 2]$. But because the complement of $(1, 2)$ is g-invariant, the g-orbit of x never enters $[1, 2]$ again. This proves that no element of $[1, 2]$ is g-periodic.

The other points shuttle between the left sub-interval $[0, 1]$ and the right sub-interval $[2, 3]$ alternately. Moreover, $g \circ g(x) = f(x)$ holds if $x \in [0, 1]$. It follows that for $0 \le x \le 1$, if $f^n(x) = x$, then $g^{2n}(x) = x$. This results in the inclusion $per(g) \supset 2per(f)$. Since every g-periodic point $y \in [2, 3]$ gives a periodic point $g(y) \in [1, 2]$ with the same g-period, there are no other g-periods. We summarize these observations as follows.

(i) If $1 \le x \le 2$, x is not g-periodic, unless x is a fixed point.

(ii) If $2 \leq x \leq 3$, x is periodic iff $x-2$ is; they have same periods.
(iii) If $0 \leq x \leq 1$, x is g-periodic iff it is f-periodic; the g-period of x is twice its f-period.

Combining these three, we conclude that

$$per(g) \;=\; \{1\} \cup \{2n : n \in per(f)\}\,.$$

Remark 6.22 If one insists that the domain of $\widetilde{f}$ also should be $[0, 1]$, (and not $[0, 3]$ as for g), then one can define

$$\widetilde{f}(x) \;=\; \frac{1}{3}g(3x) \text{ for } 0 \leq x \leq 1.$$

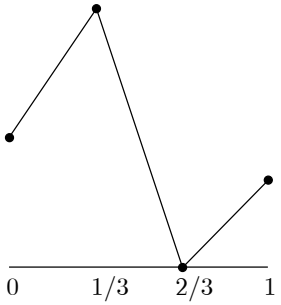
Then, $per(\widetilde{f}) = per(g)$ because $\widetilde{f}$ and g are topologically conjugate.

6.6 *Use of doubling periods in the converse of Sarkovskii's Theorem*

For the identity map, the set of periods is $\{1\}$. The main theorem of the previous section implies that for the map

$$f_1(x) \;=\; \begin{cases} x + \frac{2}{3} & \text{if } 0 \leq x \leq \frac{1}{3}; \\ 2 - 3x & \text{if } \frac{1}{3} \leq x \leq \frac{2}{3}; \\ x - \frac{2}{3} & \text{if } \frac{2}{3} \leq x \leq 1, \end{cases}$$

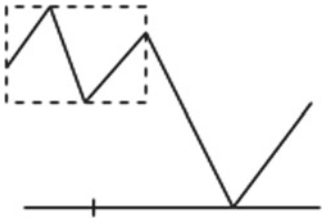
we have $per(f_1) = \{1, 2\}$.



Again, applying the same procedure, the map

$$f_2(x) \;=\; \begin{cases} f_1(x) + \frac{2}{3} & \text{if } 0 \leq x \leq \frac{1}{3}; \\ \frac{10}{3} - 5x & \text{if } \frac{1}{3} \leq x \leq \frac{2}{3}; \\ x - \frac{2}{3} & \text{if } \frac{2}{3} \leq x \leq 1, \end{cases}$$

is a piecewise linear map such that $per(f_2) = \{1, 2, 4\}$.
Proceeding like this, for each positive integer n, we have a map f_n whose set of periods is $\{1, 2, \ldots, 2^n\}$. Similarly, if f is such that $per(f)$ is the segment generated by m, then $per(\widetilde{f})$ will be the smaller segment generated by $2m$ (when m is not a power of 2).[1]



Corollary 6.23 *If 2^m forces 2^n then $m \geq n$.*

7 Proofs of Some Theorems Used

Theorem 7.1 (Baire Category theorem) *Let X be a complete metric space. Let $V_1, V_2, \ldots$ be a sequence of dense open sets in X. Then $\cap_{n=1}^{\infty} V_n$ is non-empty.*

Proof First choose $x_1 \in V_1$ and $0 < r_1 < 1$ such that the ball $B_1 = B(x_1, r_1)$ has its closure $\subset V_1$. This is possible because V_1 is nonempty (because it is dense) and open. And because if $r < s$, $\overline{B(x, r)} \subset B(x, s)$. Next choose $x_2 \in B_1 \cap V_2$ and $0 < r_2 < \frac{1}{2}$ such that the ball $B_2 = B(x_2, r_2)$ has its closure $\subset B_1 \cap V_2$. This is possible because $B_1 \cap V_2$ is nonempty (because V_2 is dense) and open. Proceeding like this, we obtain a sequence $x_1, x_2, \ldots$ in X and a sequence $r_1, r_2, \ldots$ of positive numbers with the following properties:

1. $x_i \in V_i$ for all i.
2. $0 < r_i < \frac{1}{i}$ for all i.
3. $d(x_i, x_j) < \frac{1}{i}$ for all i, if $j \geq i$.

It follows that $x_i \in B_j$ whenever $i \geq j$ and that (x_n) is a Cauchy sequence. By completeness of X, there is some $x_0 \in X$ such that $x_n \to x_0$. This $x_0 \in B_i$ for every i. Therefore $x_0 \in V_i$ for all i. □

Theorem 7.2 (Theorem 6.5) *If I and J are two closed intervals and f is a continuous real map such that $f(I) \supset J$, then there is a closed sub-interval $K \subset I$ such that $f(K) = J$.*

Proof We may assume w.l.o.g. that some element of A is less than some element of B. First we will show that if A and B are two disjoint non-empty closed subsets of

[1] The author thanks the referee for some suggestions and corrections. The author also thanks Mr. Pabitra Narayan Mandal for some useful academic discussions while preparing these notes.

I, then $\exists a \in A$ and $b \in B$ such that no element between a and b is in $A \cup B$. Take $b = \inf\{x \in B \mid x > c\}$ and $a = \sup\{x \in A \mid x < b\}$ for some $c \in A$ and $d \in B$ with $c < d$. Now it is easy to see that no element in between a and b is in $A \cup B$.

Now we use this to prove our desired result. If $J = [a, b]$, then take $A := f^{-1}(a)$ and $B := f^{-1}(b)$. Applying the previous statement we get two elements $p \in A$ and $q \in B$ such that $f([p, q]) = [a, b]$. Hence the result. □

Another Proof. The distance between the non-empty disjoint compact subsets must be attained at some $a \in A,\ b \in B$. (Because distance is a continuous function.) We claim that no element of $A \cup B$ lies between a and b. If some element p of A lies between a and b, then p will be nearer to B than a. Similarly for B. □

8 Notes & Exercises

Notes under section 1: This relates to Example (1.2.2). Obviously there is only one map satisfying $Fix(f) = \mathbb{R}$. But there are infinitely many maps satisfying $P(f) = \mathbb{R}$. It can be proved that there are only two conjugacy classes satisfying this condition. In other words, if $P(f) = \mathbb{R}$, then $f \circ f$ is identity; moreover, any two maps other than the identity map, satisfying $P(f) = \mathbb{R}$, have to be conjugate to each other.

This relates to Example (1.2.5). As the parameter μ increases from 0 to 4, the dynamics of the logistic map becomes more and more complicated. Here is a precise statement: There is a strictly increasing sequence (a_n) in the interval [2, 4] such that the following properties hold:

- If $0 < \mu < a_1$, there is an attracting fixed point and no point of period 2 or more.
- When $\mu = a_1$, a 2-cycle is born.
- If $a_1 < \mu < a_2$, there is an attracting 2-cycle, and no point of period 4 or more.
- When $\mu = a_2$, a 4-cycle is born.
- If $a_2 < \mu < a_3$, there is an attracting 4-cycle, and no point of period 8 or more.
- When $\mu = a_3$ an 8-cycle is born, and so on.
- When $\mu > a_n,\ \forall n$, all 2^k-cycles are available. And none of them is attracting.
- When μ is still greater and near 4, n-cycles are available for all n.
- When μ is 4, the system is chaotic on [0, 1].
- When $\mu > 4$, there is a chaotic subsystem topologically conjugate to the shift map.

Whatever μ be, the total number of attracting cycles is either 0 or 1.

Notes under subsection 2.2: It is worth noting that among the four ways of describing an attracting fixed point, the first uses calculus, the second uses the metric, third uses the topology and the fourth uses the dynamics. Of these, (1) and (2) are equivalent. (3) and (4) are equivalent on the real line. (2) and (3) are not equivalent. Thus there are essentially two notions of attraction.

Notes under section 3: Some other equivalent formulations of topological transitivity can be found in [13].

Notes under section 4: In [18], three piece wise linear maps are given as examples to prove a part of the main result of this section. In [2], some more examples are given. Section 4 here can be viewed as a fuller treatment of the question on hand.

Notes under subsection 5.1: This proof is modelled after the proof given by [18]. There are other proofs available in some textbooks. Our next comment pertains to Corollary 5.5. Here are a few more statements that are similar:

1. If $f^3(x) < x < f(x)$, then there is a point of period 3 between $f^3(x)$ and $f(x)$.
2. If $f^6(x) < f^5(x) < x < f(x)$, then there is point of any given period between $f^6(x)$ and $f(x)$.

See [11] for more general results of this kind.

Notes under subsection 5.2: This redundancy in Devaney's definition of chaos has been proved independently by at least three groups of researchers [4, 9, 16]; these 3 papers, appeared almost at the same time (after 3 years of publication of Devaney's book). The proof provided by us here, is along the lines of that given in [4]. In [16], only compact metric spaces are considered; and " dense orbits imply SDIC" is proved. In [9] there is a deeper study of SDIC that includes this theorem.

Notes under subsection 6.2: Here is another way to describe the Sarkovskii order on $\mathbb{N}$: There is a natural bijection $(k, l) \to 2^l(2k + 1)$ from $\mathbb{N} \times \mathbb{N}_0$ to $A := \mathbb{N} - \{\text{powers of } 2\}$. Use this to transfer the lexicographic ordering on $\mathbb{N} \times \mathbb{N}_0$ to A. After this, write the powers of 2 in the decreasing order.

Li and Yorke [12] proved that 3 forces every positive integer; that is, every interval map admitting a 3-cycle admits an n-cycle for all $n \in \mathbb{N}$. It was later observed that this is a tiny part of Theorem 6.9 that was published earlier (in [17]). Still [12] remains valuable for the notion of scrambled sets introduced therein. Many proofs are available for Theorem 6.9. The most transparent one is the graph-theoretic proof outlined in Sect. 6.

Rotation by $\frac{2\pi}{3}$ is a map from the unit circle $\mathbb{S}^1$ to itself, that admits a 3-cycle, but no other cycle. This shows that an analogue of Sarkovskii's theorem is not available on the circle $\mathbb{S}^1$.

Among the various generalizations of Sarkovskii's theorem, the following four deserve a mention here:

1. Which subsets of $\mathbb{N}$ arise as the set of all periods of continuous self maps of $\mathbb{R}$? A complete answer to this question is available as a corollary of Theorem 6.9. These sets are $\varnothing$, $\{1\}$, $\{1, 2\}$, $\{1, 2, 4\}$ etc. More precisely, a subset A of $\mathbb{N}$ arises as the set of all periods of a real map if and only if it contains all later elements of its members. (Here later means the numbers that come later in the Sarkovskii order). This formulation of the theorem leads to a natural question: For space X other than $\mathbb{R}$ or I, answer the same question: Which subsets of $\mathbb{N}$ arise as the set of all periods of continuous self maps of the space X? Complete answers are available for some spaces X. For example, when X is the unit circle, see [5] for a full answer. See [15] for further study.

2. For a real map, an n-cycle can have various order patterns. Sarkovskii's theorem is concerned with the lengths of the cycles and not with these patterns. It is natural to ask which patterns of the cycle force which others? A complete answer to this question is available in [3].
3. There are several other orbit patterns that are not cycles; they may be finite or countably infinite; there are uncountably many of them. It is natural to ask which of them forces which others. This forcing relation may not be a partial order.
4. A finite invariant set yields a finite subsystem, that is the union of finitely many orbits. It is specified by its order pattern. To start with, we may assume that f is a bijection on this subsystem. (Each cyclic pattern becomes a particular instance). We can ask which finite patterns force which others? (This is a larger question than the second question). Some preliminary investigation made in [1] indicates that this study is both challenging and rewarding.

Exercise 8.1 Find the following sets for the standard tent map (for definition see Example 1.6 or Sect. 3.2)

(i) set of all fixed points,
(ii) set of all periodic points,
(iii) set of all eventually periodic points,
(iv) set of non-wondering points.

Exercise 8.2 In any dynamical system (X, f), prove that f^m and f^n have the same set of periodic points for any $m, n \in \mathbb{N}$.

Exercise 8.3 Let $f : [0, 1] \longrightarrow [0, 1]$ be continuous map such that $f \times f$ is transitive. Prove that for every opene $U \subset [0, 1]$, $\exists x, y \in U$ and $n \in \mathbb{N}$ such that $|f^n(x) - f^n(y)| > \frac{1}{2}$.

Exercise 8.4 Let X be a Hausdorff space. Suppose f is not transitive but there exists a dense orbit. Then show that X has an isolated point.

Exercise 8.5 For all $\mu > 0$, define $q_\mu : \mathbb{R} \longrightarrow \mathbb{R}$ by $q_\mu(x) := \mu x(1 - x)$.

(i) Find all values of μ such that q_μ maps $[0, 1]$ into $[0, 1]$.
(ii) Prove that $\mu = 4$ is the only value such that $q_\mu : [0, 1] \longrightarrow [0, 1]$ is onto,
(iii) What happens when $\mu > 4$?
(iv) Find all value s of μ such that q_μ admits an attracting periodic point of period 2.

Exercise 8.6 Let X be a Hausdorff space containing infinitely many elements and $f : X \longrightarrow X$ be a continuous function. Define $N_f(U, V) := \{n \in \mathbb{N} : f^n(U) \cap V \neq \varnothing\}$. Show that

(i) for all opene sets $U, V \subset X$, $N_f(U, V) \neq \varnothing$ iff $N_f(U, V)$ is infinite.
(ii) $N_f(U, V)$ is syndetic for all opene sets $U, V \subset X$ if and only if f is minimal.
(iii) $N_f(\{x\}, V) - N_f(\{x\}, U) \subset N_f(U, V)$ for all opene sets $U, V \subset X$.

Definition 8.7 1. A dynamical system (X, f) is called minimal if X does not contain any non-empty, proper, closed f-invariant subset.
2. A syndetic set is a subset of the natural numbers, having the property of "bounded gaps", i.e., if we write the set in natural order as $\{n_k \mid k \in \mathbb{N}\}$ then $\sup_{k\in\mathbb{N}}(n_{k+1} - n_k) < \infty$.]

Exercise 8.8 Let (X, f) and (Y, g) be topologically conjugate to each other where ϕ is a topological conjugacy between them. Then

(i) both have the same number of fixed points,
(ii) ϕ takes the f-trajectory of x to the g-trajectory of $\phi(x)$,
(iii) ϕ takes f-periodic points to g-periodic points,
(iv) if x is a periodic point of f-period n, then $\phi(x)$ is a periodic point of g-period n,
(v) f is topologically transitive if and only if g is,
(vi) x is attracting fixed point for (X, f) if and only if $\phi(x)$ is attracting fixed point for (Y, g).

Exercise 8.9 Prove that the tent map and the logistic map are topologically conjugate to each other.

Exercise 8.10 This concerns Theorem 2.6. For the implication (4) $\Longrightarrow$ (5) there, the proof is completed as follows.

(a) If f is a continuously differentiable function and if p is a limit point of $Fix(f)$, then prove that $f'(0) = 1$. (Hint: Use mean value theorem).
(b) Let $g = f \circ f$. Then $g'(p) > 1$ iff $|f'(p)| > 1$.
(c) If there is no x other than p such that $g(x) = p$, then the argument in the proof of Theorem 2.6 leads to the conclusion $|f'(p)| \leq 1$. In the other case (where $g(x) = p$ for some $x \neq p$, assume w.l.o.g., $x > p$), prove that there is a fixed point strictly between p and x. Take the smallest of these fixed points and name it as q. Consider two cases.

   (i) If $g([p, q]) = [p, q]$, then prove that for every y strictly between p and q, the trajectory of y increases to q.
   (ii) If $g([p, q]) \supset [p, q]$, then prove that there is a sequence (a_n) such that $g(a_1) = q$ and $g(a_{n+1}) = a_n$ for all n. Prove that this sequence is strictly decreasing and that its limit has to be a fixed point, and hence p. Use this to arrive at a contradiction to (4).

Exercise 8.11 Let $f : X \longrightarrow X$ where X is a locally compact metric space such that $f(p) = p$. Prove that $\bigcap_{n=1}^{\infty} f^n(V) = \{p\}$ for some f-shrinking neighbourhood V of p if and only if f is a local contraction at p with respect to a finer metric with same base at p. (Outline: Define

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ d(p, x) + d(p, y) & \text{if } x \neq y; \end{cases}$$

where

$$d(p,x) \;=\; \begin{cases} \frac{1}{2^n} & \text{if } x \in f^n(V) \text{ for largest such } n \\ 1 & \text{if } x \notin V \\ 0 & \text{if } x \in f^n(V) \text{ for all } n. \end{cases}$$

Check that it is a metric and with respect to this metric f is a contraction. To prove that d is finer on X, first observe that every singleton except $\{p\}$ is open. Now take an open set containing p with respect to old metric, say W. Since $f^n(V)$ is f-shrinking to a point p, then exists some $n \in \mathbb{N}$ such that $f^n(V) \subset W$).

Exercise 8.12 Find the basin of attraction of the 2-cycle $\{1, 2\}$ for the real map $f(x) = x^2 - 4x + 5$.

Exercise 8.13 Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be strictly increasing with a discrete set of fixed points. Let

$$F_a \;=\; \{\text{attracting fixed points of } f\} \quad \text{and} \quad F_r \;=\; \{\text{repelling fixed points of } f\}.$$

Then show that F_a and F_r are intertwined as follows: between any two points of F_a there is a point of F_r; and dually.

Exercise 8.14 Prove that even for a first countable space, $R(f)$ need not be a G_δ set. (Hint: Take the space $X = [1, \omega_1)$ where ω_1 denotes the first uncountable ordinal number. Now this space with respect to order topology is first countable. Define the function

$$f(x) \;=\; \begin{cases} y & \text{if } x = y + 1 \\ x & \text{otherwise.} \end{cases}$$

Observe that f is a continuous function and $R(f)$ is the set of all limit points in this space. Every open set containing $R(f)$ has a finite complement (because the space is sequentially compact). Hence $R(f)$ is not a G_δ set.)

Exercise 8.15 This is about the converse of Theorem 1.10.

(a) Show that every closed set can be realized as the set of all fixed points of a real map. (Hint: Take $f(x) = x + d(x, F)$ where F is the given closed set.)
(b) Is this true for interval map?
(c) Show that this is not true in a general metric space.
(d) Does every F_σ-set in $\mathbb{R}$ arise as the set of all periodic points of a real map? The answer is no. The set $\mathbb{R} - \{0\}$ does not arise in this way, even though it is a F_σ-set. This leads us to the question: Which F_σ-subsets of $\mathbb{R}$ arise as $P(f)$ for some real map f? This question remains open.

References

1. Archana, M. (2018). Forcing relation and conjugacy classification using periodic points of interval maps. Ph.D. Thesis. University of Hyderabad.
2. Assaf, D., & Gadbois, S. (1992). Definition of chaos. *The American Mathematical Monthly, 99*, 865.
3. Baldwin, S. (1987). Generalizations of a theorem of Sarkovskii on orbits of continuous real-valued functions. *Discrete Mathematics, 67*, 111–127.
4. Banks, J., Brooks, J., Cairns, G., Davis, G., & Stacey, P. (1992). On Devaney's definition of chaos. *The American Mathematical Monthly, 99*, 332–334.
5. Block, L. S., & Coppel, W. A. (1992). Dynamics in one dimension. *Lecture notes in mathematics*. Berlin, Heidelberg: Springer.
6. Brin, M., & Stuck, G. (2002). *Introduction to dynamical systems*. Cambridge: Cambridge University Press.
7. Devaney, R. L. (1989). *An introduction to chaotic dynamical systems*. Addison-Wesley.
8. Elaydi, S. N. (2011). *Discrete chaos* (2nd ed.). Chapman and Hall/CRC Press.
9. Glasner, E., & Weiss, B. (1993). Sensitive dependence on initial conditions. *Nonlinearity, 6*, 1067–1075.
10. Holmgren, R. (1996). *A first course in discrete dynamical systems*. New York: Springer.
11. Li, T.-Y., Misiurewicz, M., Pianigiani, G., & Yorke, J. A. (1982). No division implies chaos. *Transactions of the American Mathematical Society, 273*, 191–199.
12. Li, T.-Y., & Yorke, J. A. (1975). Period three implies chaos. *The American Mathematical Monthly, 82*, 985–992.
13. Nagar, A., & Kannan, V. (2003). Topological transitivity for discrete dynamical systems. *Applicable Mathematics in the Golden Age* (pp. 534–584). Narosa Publications.
14. Pollicott, M., & Yuri, M. (1998). *Dynamical systems and ergodic theory*. Cambridge: Cambridge University Press.
15. Saradhi, P. V. S. P. (1997). Sets of periods of continuous self maps on some metric spaces. Ph.D. Thesis. University of Hyderabad.
16. Silverman, S. (1992). On maps with dense orbits and the definition of chaos. *Rocky Mountain Journal of Mathematics, 22*, 353–375.
17. Sarkovsky, O. M. (1964). Co-existence of the cycles of a continuous mapping of the line into itself. *Ukrainian Mathematical Journal, 16*, 61–71.
18. Vellekoop, M., & Berglund, R. (1994). On intervals, transitivity = chaos. *The American Mathematical Monthly, 101*, 353–355.

Topological Dynamics

Anima Nagar and C. R. E. Raja

Given a map $f : X \longrightarrow X$, we would like to know the asymptotic behaviour of

$$x, f(x), f^2(x), \ldots, f^n(x), \ldots$$

where $f^n(x)$ is the position of x at time n. Such a sequence is called the *trajectory* of x. This chapter comprises of the study of trajectories of all $x \in X$.

1 *G*-Spaces

One of the motivation for such a study is the following **fine dining** problem [16]:

Once upon a time lobsters were so abundant in New England waters that they were poor man's food. It even happened that prisoners in Maine rioted to demand to be fed something other than lobsters for a change. Nowadays, the haul is less abundant and lobsters have become associated with fine dining. One (optimistic?) model for the declining yields stipulates that the catch in any given year should turn out to be the average of the catches of the previous two years.

Using a_n for the number of lobsters caught in the year n, we can express this model by a simple recursion relation:

A. Nagar (✉)
Indian Institute of Technology Delhi (IITD), Delhi, India
e-mail: anima@maths.iitd.ac.in

C. R. E. Raja
Indian Statistical Institute - Bengaluru (ISI-B), Bengaluru, India
e-mail: creraja@isibang.ac.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_2

$$a_{n+1} = \frac{a_{n-1} + a_n}{2}. \tag{1}$$

As initial values, one can take the Maine harvests of 1996 and 1997 which were 16, 435 and 20, 871 (metric) tons, respectively. One can see from the recursion that all future yields should be between the two initial data. Indeed, 1997 was a record year. In fact, $a_n = x + (1 - \frac{1}{2^n})y$ gives an explicit formula for future yields (see [16] for further details and many more such examples).

Definition 1.1 Let G be a topological group and X be a (Hausdorff) topological space. We say that G acts on X or X is a *G-space* if there is a continuous map $\phi : G \times X \longrightarrow X$ such that

1. $\phi(e, x) = x$ for all $x \in X$;
2. $\phi(gh, x) = \phi(g, \phi(h, x))$ for all $g, h \in G$ and for all $x \in X$.

Here, (X, G, ϕ) is called the *transformation group*, where X is the *phase space*, G the *acting group* and the *action* ϕ gives the homeomorphism $\phi^g : X \longrightarrow X$ defined as $\phi^g(x) = \phi(g, x)$. The pair (X, G) is called a *dynamical system* or a *flow*.

Notation 1.2 For brevity, $\phi(g, x)$ will be written as gx.

Notation 1.3 Given subsets $A, B \subset G$ and $E \subset X$, we define

1. $AB = \{gh \mid g \in A, \quad h \in B\}$,
2. $A^{-1} = \{x^{-1} \mid x \in A\}$,
3. $AE = \{gx \mid g \in A, x \in E\}$—in case $E = \{x\}$, we write Ax instead of $A\{x\}$ and Ax is called the *orbit* of x under A.
4. $A_E = \{g \in G \mid gE = E\}$—in case $E = \{x\}$, we A_x instead of $A_{\{x\}}$ and A_x is called the *stabilizer* of x in A.

Remark 1.4 It is easy to see that there is a one-one correspondence between the orbit Gx and the coset space G/G_x but this need not be a homeomorphism.

Definition 1.5 For a subset $T \subset G$, a subset $E \subset X$ is called *T-invariant* if $TE \subset E$. A subset E of X is called invariant if E is G-invariant.

Definition 1.6 Let X and Y be G-spaces. A continuous map $\pi : X \longrightarrow Y$ is called a *G-map* if $\pi(gx) = g\pi(x)$ for all $x \in X$ and $g \in G$. In addition if π is a homeomorphism, we say that X and Y are homeomorphic as G-spaces.

We now look at some examples.

Example 1.7 (*Linear dynamics*) Let $X = V$ be a (finite-dimensional) vector space over a local field such as $\mathbb{R}$, the real field or $\mathbb{Q}_p$, the p-adic field. Let $G \subset GL(V)$ be the group of linear transformations on V. Then for the natural action of G on V, V is a G-space—this action of G on V is called the linear action and the corresponding dynamics is known as linear dynamics.

Example 1.8 (*Projective dynamics*) Let V be a (finite-dimensional) vector space over a local field such as $\mathbb{R}$, the real field or $\mathbb{Q}_p$. Let $\mathbb{P}(V)$ be the set of all one-dimensional subspaces of V. For a nonzero vector $v \in V$, let $\langle v \rangle$ be the one-dimensional subspace spanned by v. Thus, there is a surjective map $\pi : V \setminus \{0\} \longrightarrow \mathbb{P}(V)$ given by $\pi(v) = \langle v \rangle$. We equip $\mathbb{P}(V)$ with the smallest topology for which π is continuous. It can be seen that $\mathbb{P}(V)$ is a compact metric space.

We now define an action of $GL(V)$ on $\mathbb{P}(V)$: For $\alpha \in GL(V)$ and $\langle v \rangle \in \mathbb{P}(V)$, we define

$$\alpha\left(\langle v \rangle\right) = \langle \alpha(v) \rangle.$$

It can be verified that this defines an action of $GL(V)$ on $\mathbb{P}(V)$ and this dynamics is known as projective dynamics. Projective dynamics has proved to be fruitful in various branches of mathematics, varying from unitary representation to probability even in modern studies.

In general, one could consider Grassmannians which is the set of all r-dimensional subspaces.

Example 1.9 (*Shift map*) Let Λ be a finite set and define $X = \Lambda^{\mathbb{Z}}$. Equip X with the product topology. Then X is a compact metric space. The following is a metric on X:

$$d(u, v) = \inf \left\{ \frac{1}{k+1} \mid u(n) = v(n) \text{ for } |n| < k \right\}.$$

Define $S : X \longrightarrow X$ by $S(w)(n) = w(n+1)$. Then, S is a homeomorphism of X— specifying a homeomorphism gives rise to a $\mathbb{Z}$-action and specifying d-commuting homeomorphisms give rise to $\mathbb{Z}^d$-action.

Example 1.10 (*Algebraic dynamics*) Let X be a topological group and G be a group of automorphisms of X. Then. the canonical action of G on X is defined by

$$\phi(\alpha, x) = \alpha(x)$$

for all $\alpha \in G$ and $x \in X$.

An interesting example is the case when $X = \mathbb{T}^n$ is the n-dimensional torus (realised as the quotient group $\mathbb{R}^n/\mathbb{Z}^n$ or direct product of n-copies of the circle $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$). In this case, the group of automorphisms is identified with $GL_n(\mathbb{Z})$.

In the next set of remarks we assume that X is a G-space. The following can be taken as exercises:

1. Let $E \subset X$. Then E is G-invariant, that is $GE \subset E$ iff $gE \subset E$ for all $g \in G$.
2. X and $\varnothing$ are G-invariant.
3. If E is G-invariant, then $X \setminus E$, $\overline{E}$, E^o are also G-invariant.
4. If E and F are G-invariant, then $E \cap F$ is also G-invariant.

5. An arbitrary union and arbitrary intersection of G-invariant sets are also G-invariant.
6. Let $T \subset G$ and $E \subset X$. Then E is T-invariant iff $X \setminus E$ is T^{-1}-invariant.
7. Gx is the smallest G-invariant set containing x.
8. $y \in Gx$ if and only if $Gy = Gx$.
9. $X = \cup Gx$—disjoint union.
10. $\overline{Gx}$ is called the orbit closure of x under G and is the smallest closed G-invariant set containing x.
11. $y \in \overline{Gx}$ implies $\overline{Gy} \subset \overline{Gx}$.

Definition 1.11 Let $\{(X_i, G_i, \phi_i)\}$ be a collection of dynamical systems. The *product dynamical system* denoted by $\prod_i (G_i, X_i, \phi_i)$ is the dynamical system (G, X, ϕ) where $G = \prod_i G_i$, $X = \prod_i X_i$ and $\phi : G \times X \longrightarrow X$ is defined by $\phi(g, x) = (\phi_i(g_i, x_i))$ for any $g = (g_i) \in G$ and $x = (x_i) \in X$.

Definition 1.12 Let X_i be a collection of G-spaces. Then the *product of* G-spaces (G, X, ϕ), where $X = \prod_i X_i$ with the G-action, is defined by $gx = (gx_i)$ for any $g \in G$ and $x = (x_i) \in X$.

The most intensively studied case is when the acting group $G = \mathbb{Z}$. In that case $\phi^1 = f$ gives a generating homeomorphism on X, i.e., $f(x) = \phi(1, x)$ giving iterations $f^n(x) = \phi(n, x) = nx$. We call the system (X, f) a *cascade*.

Many times, we are just interested in a semigroup S and we have a semigroup action (S, X, ψ). We study the *semi-cascade* $(\mathbb{N}, X, \psi)$ or (X, f) where $f : X \longrightarrow X$ is a continuous mapping.

Here $\mathcal{O}(x) = \{f^n(x) : n \in \mathbb{N}\}$ is called the *orbit* of the point x.

Definition 1.13 For a cascade or semi-cascade (X, f), $x_0 \in X$ is called a *fixed point* if $f(x_0) = x_0$. And $y_0 \in X$ is called a *periodic point* if there exists $n \in \mathbb{N}$ such that $f^n(y_0) = y_0$. The smallest such n is called the *period* of y_0.

Definition 1.14 The *ω-limit set* of a point $x \in X$ under f, denoted as $\omega(x)$, is the set of all limit points of $\{f^n(x) : n \in \mathbb{Z}(\mathbb{N})\}$, and is a non empty closed f-invariant set.

Definition 1.15 A point $x \in X$ is said to be *non-wandering* if for every neighbourhood U of x there is a $n \in \mathbb{N}$ such that $f^n(U) \cap U \neq \varnothing$. The set of all non-wandering points of f is denoted as $\Omega(f)$.

Definition 1.16 For a flow (X, G), a point $x \in X$ is called *recurrent* when $x \in \omega(x)$, i.e., if the orbit of x returns to its neighbourhood infinitely often. Usually this 'infinitely often' is described in terms of some *admissible set*. These admissible sets are either in the form of *extensive sets* as considered by Gottschalk and Hedlund or more extensively in the form of *(Furstenberg) Families*—mimicking the aspects of recurrence studied by Furstenberg.

For a cascade or semi-cascade (X, f), a point $x \in X$ is called *recurrent* when $x \in \omega(x)$, i.e., for every open set U containing x there exist $j \in \mathbb{N}$ such that $f^j(x) \in U$, i.e., there exists a sequence $n_k \nearrow \infty$ such that $f^{n_k}(x) \to x$ i.e., the set $N(x, U) = \{n \in \mathbb{N} : f^n(x) \in U\}$, of *return times*, is infinite for any neighbourhood U of x.

The set of all recurrent points in X is denoted as $\mathcal{R}(X)$.

2 Minimal Systems

Let (X, G) be a flow, where X is compact. The simplest dynamics that one can observe is when the system is 'minimal'.

Definition 2.1 A set $M \subset X$ is called a *minimal set* if M is closed, nonempty and invariant and M has no proper subset with these properties, i.e., if $N \subseteq M$ is closed and invariant, then $N \equiv M$ or $N \equiv \varnothing$.

Proposition 2.2 *M is minimal if and only if it is the orbit closure of each of its points, i.e., $\forall x \in X,\ M = \overline{Gx}$.*

Proof Let M be minimal and $x \in M$. Then $\overline{Gx} \subseteq M$ is nonempty, closed and invariant. Hence, $\overline{Gx} = M$.

Conversely, if $\overline{Gx} = M, \forall x \in M$, but M is not minimal, then there exists $N \subset M$ which is closed and invariant and so for $x \in N$, we have $\overline{Gx} \subseteq N$. Thus $M = N$ and so M is minimal. □

Definition 2.3 If $X = \overline{Gx},\ \forall x \in X$, then the flow (X, G) is called a *minimal flow*.

Proposition 2.4 *If M_1 and M_2 are minimal subsets of X for any flow (X, G) then either $M_1 = M_2$ or $M_1 \cap M_2 = \varnothing$.*

Proposition 2.5 *Let (X, G) be a flow. Then X contains a minimal set.*

Proof Let $\mathcal{M}$ denote the set of all nonempty closed, invariant subsets of X. Then, $X \in \mathcal{M}$ and so $\mathcal{M} \neq \varnothing$. Also $\mathcal{M}$ is a partially ordered set. Consider a chain $\{M_\alpha\}$ in $\mathcal{M}$. Then $M^* = \cap M_\alpha$ is nonempty and is also closed and invariant, and so $M^* \in \mathcal{M}$. Thus every chain in $\mathcal{M}$ is bounded below and so by Zorn's lemma, $\mathcal{M}$ contains a minimal element, which is the minimal subset of X. □

For cascades, fixed points and periodic points are trivial minimal subsets.

Definition 2.6 $A \subset G$ is called *syndetic* if there is a compact $K \subset G$ such that $G = KA = \{ka : k \in K \text{ and } a \in A\}$.

$A \subset \mathbb{Z}(\mathbb{N})$ is called syndetic if it is relatively dense i.e. does not contain arbitrarily large gaps or has bounded gaps.

Definition 2.7 For a flow (X, G), a point $x \in X$ is called an *almost periodic point* if for every neighbourhood U of x, there is a syndetic $A \subset G$ such that $Ax \subset U$.

For a cascade $(X, f),\ x \in X$ is *almost periodic* if for any neighbourhood $U \ni x$, the set $N(x, U) = \{n \in \mathbb{N} : f^n(x) \in U\}$ is syndetic.

Theorem 2.8 *For a flow (X, G), a point $x \in X$ is an almost periodic point if and only if $\overline{Gx}$ is minimal.*

Proof Let $A = \overline{Gx}$ be minimal and let U be a neighbourhood of x. We note that we must have $A \subset GU$, else $A \setminus GU$ becomes a closed, invariant subset of the minimal A. Since A is compact, there are finitely many $g_1, \ldots, g_n \in G$ such that $A = \bigcup_{i=1}^{n} g_i U$. Now for any $h \in G$, there exists an i, $1 \leq i \leq n$, such that $hx \in g_i U$ and so $g_i^{-1} hx \in U$. Let $K = \{g : gx \in U\}$ then $g_i^{-1} h \in K$, i.e., $h \in g_i K \subset \{g_1, \ldots, g_n\}K$. Thus $G = \{g_1, \ldots, g_n\}K$ and so K is syndetic implying that x is almost periodic.

Conversely, suppose x is almost periodic but $\overline{Gx}$ is not minimal. Then being compact $\overline{Gx}$ contains a minimal set M. Clearly $x \notin M$. Let U and V be disjoint open sets with $x \in U$ and $M \subset V$. Let $H \subset G$ be compact. Let $HW \subset V$ for some open $W \supset M$.

Since $M \subset \overline{Gx}$, there is a $h \in G$ such that $hx \in W$. Then $Hhx \subset HW \subset V$. So $Hhx \cap U = \varnothing$. So if $K = \{g : gx \in U\}$ then $G \neq HK$. H being arbitrary, this implies that K is not syndetic and so x is not almost periodic. This contradiction proves the assertion. □

We look at some examples of minimal flows.

Example 2.9 Let $\mathbb{T} = \{z : |z| = 1\}$ be the unit circle in $\mathbb{C}$. Let $\alpha = e^{i2\pi\theta} \in \mathbb{T}$ with θ irrational. Note that $\alpha^n \neq 1\ \forall n \in \mathbb{N}$.

Define $\tau : \mathbb{T} \to \mathbb{T}$ as $\tau z = \alpha z$. We will show that $\mathcal{O}(1) = \{\alpha^n : n \in \mathbb{N}\}$ is dense in $\mathbb{T}$. This proves that $\mathcal{O}(z) = \{\alpha^n z : n \in \mathbb{N}\}$ is dense in $\mathbb{T}$, proving the minimality of $(\mathbb{T}, \tau)$.

Let $\beta \in \overline{\mathcal{O}(1)}$ and let $\epsilon > 0$. Then there exists $n, k > 0$ such that $|\alpha^n - \beta| < \frac{\epsilon}{2}$ and $|\alpha^{n+k} - \beta| < \frac{\epsilon}{2}$. Thus $|\alpha^{n+k} - \alpha^n| < \epsilon$. Since the map $z \to \alpha^k z$ is an isometry,

$$\cdots |\alpha^{n+3k} - \alpha^{n+2k}| = |\alpha^{n+2k} - \alpha^{n+k}| = |\alpha^{n+k} - \alpha^n| < \epsilon.$$

For some $m \in \mathbb{N}$, the points $\alpha^n, \alpha^{n+k}, \ldots, \alpha^{n+mk}$ wind around the circle. Thus $\mathcal{O}(1)$ is $\epsilon-$dense in $\mathbb{T}$, giving our requirement.

Example 2.10 We now recall the example (1.9), where $\Lambda = \{0, 1\}$ and $X = \Lambda^{\mathbb{Z}}$. We consider the shift map $S : X \longrightarrow X$.

To obtain a minimal subset of X, it is enough to construct an almost periodic point $p \in X$ since then $\overline{\mathcal{O}(p)}$ will be minimal. This can precisely be done by the rich theory of Jewett-Krieger constructions developed independently by R. Jewett and W. Krieger during 1969–1971.

Here we look into a classical construction due to Marston Morse, originally developed by Axel Thue who used it in the study of combinatorics on words. This construction is done using substitution: $0 \to 01,\ 1 \to 10$. Hence,

$$0 \to 01 \to 0110 \to 01101001 \to 0110100110010110 \to \cdots .$$

This will finally converge to some $x \in \{0, 1\}^{\mathbb{N}}$. This construction indicates that every finite word in x occurs syndetically often. Extend x to $p \in X$ by defining

$$p(n) = \begin{cases} x(n), & n \geq 1; \\ x(-n-1), & n < 0. \end{cases}$$

Every word in p occurs syndetically and p is symmetric at the mid point, and so p is almost periodic. $\overline{\mathcal{O}(p)} \subset X$ is a minimal subset of X.

Exercise 2.11 For $G-$spaces X and Y, let $\pi : X \longrightarrow Y$ be a $G-$map:

1. If $X_0 \subset X$ is minimal, then $\pi(X_0) = Y_0 \subset Y$ is minimal.
2. If X is minimal and both π and ψ are $G-$maps that agree on a point, then $\pi = \psi$.
3. If $X_0 \subset X$ and $Y_0 \subset Y$ are minimal such that $\pi(X_0) \cap Y_0 \neq \varnothing$, then $\pi(X_0) = Y_0$.
4. If X is minimal then the only closed, invariant subsets of X are $\varnothing$ and X.

Recall the notion of recurrence in X. We see that almost periodicity is a very strong form of recurrence. We recall a strong result guaranteeing recurrence for cascades.

Theorem 2.12 (Birkhoff Recurrence Theorem) *For a cascade (X, f), there exists $x \in X$ such that $f^{n_i}x \to x$ for some sequence $\{n_i\}$ in $\mathbb{N}$, i.e., $\mathcal{R}(X) \neq \varnothing$.*

A simple proof for this is to first apply Zorn's lemma to show that every cascade admits a minimal subset and then use the existence of an almost periodic point for this recurrent point.

3 Multiple Recurrence and Van Der Waerden's Theorem

A good number of results in combinatorial number theory have the following general form: *For any finite partition of the natural numbers $\mathbb{N}$ into classes $C_1, C_2, \ldots, C_r$, at least one of the classes possesses property P*. For example, if P is the property that a subset contains arithmetic progressions of arbitrary finite length, then the aforementioned becomes the well-known theorem of van der Waerden. We now discuss the topological dynamics proof of van der Waerden's Theorem due to Furstenberg and Weiss, [13]—see [12, 13] for more related results. A useful tool is the following multiple recurrence theorem.

Theorem 3.1 (Multiple recurrence) *Let X be a compact metric space and $T_1, T_2, \ldots, T_p$ be commuting homeomorphisms of X. Then there exists a point $x \in X$ and a sequence (k_n) such that $T_i^{k_n}(x) \to x$ simultaneously for $i = 1, 2, \ldots, p$.*

Exercise 3.2 (*Multiple recurrence*)

1. Let $T_1, T_2 : \mathbb{T}^2 \longrightarrow \mathbb{T}^2$ be given by

$$T_1(x, y) = (xy, y) \quad \text{and} \quad T_2(x, y) = (x, xy).$$

Find the multiple recurrence points for T_1 and T_2.

2. Consider the projective linear transformations T_α and T_β on the projective line defined by the linear maps

$$\alpha = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Are there any multiple recurrent points for T_α and T_β?

We first derive van der Waerden's Theorem from Multiple recurrence.

Theorem 3.3 (van der Waerden) *For any finite partition $\mathbb{N} = C_1 \cup C_2 \cup \cdots \cup C_p$, there is a C_i containing arithmetic progression of arbitrary finite length, that is, for each $n \geq 1$, there is a $x_n, y_n \in C_i$ such that $y_n \neq 0$ and $x_n + ky_n \in C_i$ for all $1 \leq k \leq n$.*

Exercise 3.4 (*Infinite multiple*)

1. Give an example to show that Multiple recurrence theorem is not true for infinite number of commuting transformations.
2. Give an example of a finite partition $\mathbb{Z}$ such that none of the sets contains an infinite arithmetic progression.

Proof Let $\mathbb{N} = C_1 \cup C_2 \cup \cdots \cup C_p$ be a partition. Take $A = \{1, 2, \ldots, p\}$ and $\Omega = A^{\mathbb{Z}}$, the space of A-valued sequences. Define the shift $S : \Omega \to \Omega$ given by $Sw(n) = w(n+1)$. We endow Ω with the metric

$$d(u, v) = \inf \left\{ \frac{1}{k+1} \mid u(n) = v(n) \text{ for } |n| < k \right\}.$$

We have already seen that Ω is compact and S is a homeomorphism of Ω.

Define $\psi \in \Omega$ by $\psi(k) = j$ if $k \in C_j$, otherwise $\psi(k) = 1$.

Let X be the set of limit points of the sequence $(S^n\psi)_{n\geq 1}$. Then X is S-invariant. Now take $T_i = S^i$ for $i = 1, \ldots, p$. Applying multiple recurrence theorem to $(X, T_1, \ldots, T_p)$, we get $\eta \in X$ and $n \in \mathbb{N}$ such that

$$d\left(T^i\eta,\ \eta\right) \ < \ \frac{1}{2} \ \forall i.$$

This implies by evaluating at $k = 0$, that $\eta(0) = \eta(in)$ for all $i = 1, \ldots, p$.

Choose $m > 0$ so that $d(S^m(\psi), \eta) < \dfrac{1}{pn+1}$. We then have $\psi(m) = \psi(m + in)$ for all $i = 1, \ldots, p$. This implies that $m + in \in C_{\psi(m)}$, for all $i = 1, \ldots, p$, that $C_{\psi(m)}$ contains an arithmetic progression of length $p + 1$. □

We now prove multiple recurrence theorem in a series of lemmas.

Lemma 3.5 *A dynamical system (X, G) is minimal if and only if for any $\epsilon > 0$ there exists a finite set of transformations $S_1, S_2, \ldots, S_k$ of G such that for any $x, y \in X$, $\min_i d(S_i x, y) < \epsilon$.*

Proof If V is any open subset of X, $\cup_{S\in G} S^{-1}V$ is an open G-invariant set which by minimality is all of X. Since X is compact, a finite subcovering covers X. Letting V range over a finite cover of X by sets of diameter less than ϵ, we obtain the condition of the lemma. The converse is clear. □

Lemma 3.6 *Let (X, T) be a dynamical system with X being a compact metric space and A, a closed subset of X such that*

1. *there is group G that acts on X commuting with T such that A is G-invariant and minimal;*
2. *for each $\epsilon > 0$ there exist $x, y \in A$ and $n \geq 1$ with $d(T^n x, y) < \epsilon$.*

Then for each $\epsilon > 0$ there is a $z \in A$ and $n \geq 1$ such that $d(T^n z, z) < \epsilon$.

Proof due to R. Bowen. Let $\epsilon > 0$ be arbitrary. We now inductively construct sequences (z_k) in A, (n_r) in $\mathbb{N}$ and $\epsilon_r < \frac{\epsilon}{2}$ such that $d(T^{n_r} z_r, z_{r-1}) < \epsilon_r$ and $d(T^{n_r} z, z_{r-1}) < \epsilon_r$ whenever $d(z, z_r) < \epsilon_{r+1}$.

Fix $z_0 \in A$. Then by (3.5), there exists $S_1, \ldots, S_N$ in G be such that

$$\min_i d(S_i x, z_0) < \frac{\epsilon}{4}, \quad \text{for any } x \in A. \tag{2}$$

Let $\delta > 0$ be such that $d(x, y) < \delta$ implies that $d(S_j x, S_j y) < \frac{\epsilon}{4}$. Now choose x_1, y_1 in A and n_1 so that $d(T^{n_1} x_1, y_1) < \delta$—possible by assumption. Then, $d(S_j T^{n_1} x_1, S_j y_1) = d(T^{n_1} S_j x_1, S_j y_1) < \frac{\epsilon}{4}$, and combining this with (2), we obtain $(T^{n_1} S_i x_1, z_0) < \frac{\epsilon}{2}$ for some i.

Take $z_1 = S_i x_0$. Then $d(T^{n_1} z_1, z_0) < \epsilon_1$ for some $\epsilon_1 < \frac{\epsilon}{2}$.

Proceeding inductively, we obtain the sequences (z_k) in A, (n_r) in $\mathbb{N}$ and $\epsilon_r < \frac{\epsilon}{2}$ such that $d(T^{n_r} z_r, z_{r-1}) < \epsilon_r$ and $d(T^{n_r} z, z_{r-1}) < \epsilon_r$ whenever $d(z, z_r) < \epsilon_{r+1}$—for the second inequality use also the continuity of T^{n_r}. Thus, we have

$$d(T^{n_j + n_{j-1} + \cdots + n_{i+1}} z_j, z_i) < \frac{\epsilon}{2} \quad \text{for } j > i. \tag{3}$$

Since A is compact, there exists $i < j$ such that $d(z_i, z_j) < \frac{\epsilon}{2}$ and hence, using (3), we get that $d(T^n z_j, z_j) < \epsilon$ for $n = n_j + n_{j-1} + \cdots + n_{i+1}$. □

Lemma 3.7 *Let X, T and A be as in (3.6). Then, there is a point $z \in A$ which is recurrent for T.*

Proof Let $F(x) = \inf_n d(T^n x, x)$. Then F is upper semicontinuous, and contains a point of continuity. Let x_0 be a point of continuity. We claim that $F(x_0) = 0$, that would prove that x_0 is recurrent for T.

If $F(x_0) > 0$, then there is an open set V containing x_0 such that $F(x) > \delta > 0$ for all $x \in V$. Then, there exists $S_1, \ldots, S_N$ in G such that $A \subset \cup_{i=1}^N S_i^{-1} V$. Let $\eta > 0$ be

such that $d(S_i x, S_i y) < \delta$ whenever $d(x, y) < \eta$ for all i. If $F(x) < \eta$, then for some n, we have $d(T^n x, x) < \eta$ and hence, $d(T^n S_i x, S_i x) < \delta$. This implies that $F(S_i x) < \delta$ and hence, $S_i x \notin V$. Thus, $F(x) \geq \eta$ for all $x \in A$. This is a contradiction to (3.6). □

Exercise 3.8 (*Dense set of recurrent points*)

1. In (3.7), prove that the set of points in A that are recurrent for T is dense in A.
2. In multiple recurrence theorem, if X is minimal, then prove that the set of multiple recurrent points is dense in X.
3. Prove by example that minimality is required to get dense set of recurrent points.

Proof of multiple recurrence theorem. Let G be the group generated by $T_1, \ldots, T_p$. By restricting to a minimal closed invariant set of X, we may assume that X is minimal. For $p = 1$, the result follows from Birkhoff's theorem but it also follows from (3.7).

Let A denote the diagonal in X^p and $T = T_1 \times T_2 \times \cdots \times T_p$ on X^p. Note that G also acts X^p as X^p is a product of G-spaces. Then A is a G-invariant minimal closed set as X is G-minimal.

We claim that A satisfies the hypotheses of (3.6) with respect to T. Let $R_i = T_i T_p^{-1}$ for $i = 1, \ldots, p-1$. Then, by induction assumption, there exists $x \in X$ such that $R_i^{k_n} x \to x$ for all $i = 1, \ldots, p-1$.

For any $\epsilon > 0$, there is a n such that

$$d(T^n z, y) = \sum_{i<p} d(R_i^n x, x) < \epsilon, \quad \text{for } y = (x, x, \ldots, x) \text{ and } z = T_p^{-n}(y) \in A.$$

Thus, the hypothesis of (3.6) is verified and hence, the result follows from (3.7). □

4 Enveloping Semigroups

The *Enveloping Semigroup* of a flow (X, G) was introduced by Robert Ellis in [7], though it was considered in his earlier work on distality. It can be viewed as a compactification of the acting group G.

Let X be a compact, Hausdorff space. Then the set of all self maps (not necessarily continuous) on X can be identified with X^X endowed with the product topology, which is basically the space of all functions on X with the point-open topology. X^X is compact by the Tychnoff's theorem. X^X also has a semigroup structure, given by composition of functions, i.e., for $\alpha, \beta \in X^X$, $\alpha\beta(x) = \alpha(\beta(x))$ at the x^{th} coordinate in X^X.

For the flow (X, G, ϕ), the *evaluation map* ϕ is defined as $\phi(g, x) = g(x)$. Let $\widetilde{G} = \{\phi^g : g \in G\}$ where $\phi^g : X \longrightarrow X$ with $\phi^g(x) = gx$. Then $\widetilde{G} \subset X^X$. We can identify G with $\widetilde{G}$ and consider $G \subset X^X$.

Definition 4.1 *The Enveloping Semigroup* of the flow (X, G) is defined as $E = E(X) = E(X, G) = \overline{G}(= \widetilde{\overline{G}}) \subset X^X$.

We note that $E(X)$ is also compact and Hausdorff. Now G acts on X^X via composition and we have $h(\phi^g) = \phi^{hg}$ and so $G \circ G = G$ and hence $E(X)$ is invariant under G. Thus $(E(X), G)$ is also a flow i.e., $E(X)$ is also a $G-$space.

Now since $E(X) = \overline{G}$, for every $\alpha \in E(X)$, there is a net $\{g_t\}$ in G such that $g_t \to \alpha$, i.e., $\phi^{g_t} \to \alpha$ pointwise. Hence for $\beta \in E(X)$, $\beta g_t \to \beta\alpha$. Now since $\beta g_t \in E(X)$ and $E(X)$ is closed, $\beta\alpha \in E(X)$ giving $E(X)^2 = E(X) \circ E(X) \subset E(X)$. Thus $E = E(X)$ is a subsemigroup of X^X.

Remark 4.2 Though the elements of G are homeomorphisms, the elements of $E(X)$ need not even be continuous.

Proposition 4.3 *Let $E(X)$ be the enveloping semigroup for the flow (X, G). Then, $\forall x \in X$, $Ex = \overline{Gx}$.*

Proof Let $y \in \overline{Gx}$. Then, there is a net $\{g_t\}$ in G such that $g_t(x) \to y$. For this net, there is a subnet $\{g_{t_s}\}$ such that $\{g_{t_s}\} \to \alpha$ in $E(X)$. Hence $\{g_{t_s}\}(x) \to \alpha(x)$ i.e., $\{g_t(x) \to \alpha(x)\}$ and so $y = \alpha(x) \in Ex$. Thus, $\overline{Gx} \subset Ex$.

The converse follows by taking the reverse path in the above argument. □

Remark 4.4 For any $y \in \overline{Gx}$, there exists $p \in E(X)$ such that $y = p(x)$.

Definition 4.5 (X, G) is called an *equicontinuous flow* if $\{\phi^g : g \in G\}$ is an equicontinuous family on X.

If (X, G) is an equicontinuous flow, then the topology on G considered as a subset of X^X is the topology of uniform convergence and so $\overline{G} = E$ is compact and equicontinuous by Arzela-Ascoli's theorem. Working along this line, we can prove:

Theorem 4.6 *A flow is equicontinuous if and only if its enveloping semigroup is a group of homeomorphisms.*

Lemma 4.7 *Let (X, G) be a flow with G being an Abelian. Then,*

1. *for $p \in E(X)$ and $g \in G$, $gp = pg$.*
2. *if (X, G) is equicontinuous, then $E(X)$ is Abelian.*

Proof Let $p \in E(X)$. Take a net $\{g_t\}$ in G with $g_t \to p$. Then $gg_t \to gp$. But $gg_n = g_n g \to pg$ and so $gp = pg$.

Now when (X, G) is equicontinuous, then all $p \in E(X)$ are continuous and so $pq = qp$ $\forall p, q \in E(X)$. □

Proposition 4.8 *If (X, G) is equicontinuous, then the flow $(E(X), G)$ is minimal.*

Proof Let e be the identity in G, then $\overline{\mathcal{O}(e)} = \overline{\{ge : g \in X\}} = E(X)$. Also, $(E(X), G)$ is equicontinuous and hence $(E(X), G)$ must be minimal. □

Exercise 4.9 Let X and Y be $G-$spaces and $\pi : X \longrightarrow Y$ be a $G-$map. Then,

1. there exists a unique $G-$map $\theta : E(X) \longrightarrow E(Y)$.
2. if $\pi' : X \longrightarrow Y$ is another $G-$map, then both π and π' induce the same $G-$map $\theta : E(X) \longrightarrow E(Y)$.

Definition 4.10 Let I be a non-vacuous subset of $E = E(X)$. Then I is said to be a left ideal in $E(X)$ or simply an *ideal* if $EI \subset I$. I is said to be a *minimal ideal* if whenever K is a non-vacuous subset of I such that $EK \subset K$, we have $K = I$.

Lemma 4.11 *For a flow (X, G), the following holds:*

1. *Let $\varnothing \neq I \subset E(X)$. Then I is a minimal ideal if and only if the flow (I, G) is minimal.*
2. *Let I be a minimal ideal in $E(X)$. Then Ix is a minimal subset of X for all $x \in X$, where $Ix = \{px : p \in I\}$.*

Since every compact system contains a minimal subset, $E(X)$ contains minimal ideals. Minimal ideals have a rich algebraic structure. We look into the following which was first proved by K. Numakura in [20].

Theorem 4.12 *For a flow (X, G), every minimal ideal I contains an idempotent.*

Proof Let $\mathcal{M}$ be the collection of all closed subsets A of I such that $A^2 \subset A$. Since $I \in \mathcal{M}$, $\mathcal{M} \neq \varnothing$. By Zorn's lemma $\mathcal{M}$ has a minimal element say S. If $x \in S$, then $(Sx)^2 = (Sx)(Sx) \subset S(Sx) = S^2x \subset Sx$. Hence $Sx \subset \mathcal{M}$. Also $Sx \subset S^2 = S$ and hence by minimality $Sx = S$.

Let $p \in S$ be such that $px = x$ and let $\mathcal{K} = \{\alpha \in S : \alpha x = x\} \subset S$. Then $p \in \mathcal{K}$ and so $\mathcal{K} \neq \varnothing$. If $a, b \in \mathcal{K}$, then $ab(x) = a(x) = x$, and $\mathcal{K}^2 \subset \mathcal{K}$. Thus $\mathcal{K} = S$ and so $x^2 = x$. This $x \in I$ is an idempotent. □

Idempotents are important elements of enveloping semigroups.

Theorem 4.13 *Let (X, G) be a flow and I be a minimal ideal in $E(X)$. TFAE:*

(a) $x \in X$ is an almost periodic point.
(b) $ux = x$ for some idempotent $u \in I$.

Exercise 4.14 Let $I \subset E(X)$ be a minimal ideal. Then

(i) $Ip = I$, for all $p \in I$.
(ii) $up = p$, for $u, p \in I$ and u is an idempotent.
(iii) If $u \in I$ is an idempotent and $p \in I$ with $up = u$, then p is also an idempotent.
(iv) If $u \in I$ is an idempotent then Iu is a group with identity u.
(v) If $p \in I$ then there is a unique idempotent $u \in I$ with $up = p$.
(vi) If $u, v \in I$ are idempotents, with $u \neq v$, then $Iu \cap Iv = \varnothing$.

Definition 4.15 Let $x, y \in X$. Then x and y are said to be *proximal* if there exists a net $\{g_t\}$ in G with $\lim g_t(x) = \lim g_t(y)$. The *proximal relation* $P(X)$ is defined to be that subset of $X \times X$ consisting of all proximal pairs (x, y).

The relation $P(X)$ is reflexive and symmetric, but in general not transitive. We note the following result first proved in [7].

Theorem 4.16 *The following statements are equivalent.*

1. $P(X)$ is an equivalence relation on X.
2. $E(X)$ contains exactly one minimal ideal.

Enveloping semigroups are very useful in studying proximal pairs.

Theorem 4.17 *Two points $x, y \in X$ are proximal if and only if $px = py$ for some $p \in E(X)$.*

Proof Let x, y be proximal. Then, there exists a net $\{g_t\}$ in G and $z \in X$ such that $g_t x \to z$ and $g_t y \to z$. If necessary, by passing to a subnet, we get a $p \in E(X)$ such that $g_t \to p$. Then $px = z = py$.

Conversely, let $p \in E(X)$ be such that $px = py = z$. Then there exists a net $\{g_t\}$ in G such that $g_t \to p$. Hence, $g_t x \to px = z$ and $g_t y \to py = z$ implying that x, y are proximal. □

Definition 4.18 A flow (X, G) is called *distal* if it has no non-trivial proximal pairs. That is, $P(X) = \triangle = \{(x, x) : x \in X\} \subset X \times X$.

If (X, d) is a (compact) metric space, then (X, G) is distal if and only if $\inf_{g \in G} d(gx, gy) > 0$, whenever $x \neq y$.

Remark 4.19 When G is a group, all equicontinuous flows will be distal, but distal flows need not be equicontinuous.

Let $D = \{(r, \theta) : 0 \leq r \leq 1, \ 0 \leq \theta \leq 2\pi\}$, and $f : D \longrightarrow D$ defined as $f(r, \theta) = (r, \theta + r(\mathrm{mod} 2\pi))$. Then the cascade (D, f) is distal but not equicontinuous.

Theorem 4.20 (Auslander-Ellis) *For the flow (X, G), every point is proximal to an almost periodic point.*

This gives that every point is almost periodic for a distal flow.

Theorem 4.21 *(X, G) is a distal flow if and only if $E(X)$ is a group.*

Proof Let $E(X)$ be a group, and $(x, y) \in P(X)$. Then there exists a $p \in E(X)$ such that $px = py$. But this gives that $x = p^{-1}px = p^{-1}py = y$, and so (X, G) is distal.

To prove the converse, we recall some facts:

1. For a collection of flows (X_α, G), $E(\prod_\alpha X_\alpha, G) = \triangle \prod_\alpha E(X_\alpha, G)$.
2. For a collection of flows (X_α, G), $(\prod_\alpha X_\alpha, G)$ is distal if and only if each (X_α, G) is distal.

Since (X, G) is distal, the product flow (X^X, G) is also distal, so by theorem (4.20) every $\phi \in X^X$ is almost periodic. So every $p \in E(X)$ is also almost periodic.

For $p \in E(X)$, $p \in Ee = \overline{Ge}$, and since $\overline{Ge}$ is minimal, $e \in Ep = \overline{Gp}$ i.e., there exists $q \in E(X)$ such that $e = qp$. So every $p \in E(X)$ has a left inverse. It is a simple to see that q will also be a right inverse of p and so $E(X)$ is a group. □

Proximal and distal flows are important class of G-spaces. We study more details about them in the next section.

We consider an alternate definition of the enveloping semigroup. For a topological group G, the *Stone-Čech compactification* βG of G is determined upto homeomorphism as:

(a) $G \subset \beta G$, with βG compact, Hausdorff;
(b) $\overline{G} = \beta G$;
(c) if Z is a compact Hausdorff space then, any function $f : G \longrightarrow Z$ has a unique continuous extension $\hat{f} : \beta G \longrightarrow Z$.

Let G have the identity element e. G is provided with an associative binary operation: $(g, h) \longmapsto gh$, then the left multiplication $h \longmapsto gh$ is continuous for all $g \in G$ and the right multiplication $q \longmapsto qp$ is also continuous for all $p \in \beta G$. Thus, the group structure of G can be extended to the semigroup structure of βG with left and right multiplication continuous. Again, G acts on βG and so $(\beta G, G)$ is a flow.

Lemma 4.22 *Let (X, G) have a dense orbit, i.e., $X = \overline{Gx}$ for some $x \in X$. Then there exists a G-map $f : \beta G \longrightarrow X$.*

Proof Since $X = \overline{Gx}$, the map $f : G \longrightarrow X$ defined as $g \longmapsto gx$ has a unique continuous extension $\hat{f} : \beta G \longrightarrow X$ given by $p \longmapsto px$. Now $g(px) = gp(x)$ and so $\hat{f}$ is a G-map, and $X = \overline{Gx} \subset \overline{\hat{f}(G)} \subset \hat{f}(\overline{G}) = \hat{f}(\beta G)$. □

Let (X, G) be a flow. Then the identification $\psi : G \longrightarrow X^X$ has a continuous extension $\Psi : \beta G \longrightarrow X^X$,

$$\Psi(\beta G) = \Psi(\overline{G}) = \overline{\{\phi^g : g \in G\}} = E(X).$$

$E(\beta G) \cong \beta G$. In fact, the largest possibility for any $E(X, G)$ is βG.

5 Proximal and Distal

We assume that X is a metrizable G-space with metric d. Recall the following:

1. A pair of points $x, y \in X$ is proximal if

$$\inf_{g \in G} d(gx, gy) = 0$$

and if every pair in X is proximal, we say that the action of G on X is proximal or G is proximal on X.

2. A point in $x \in X$ is called distal if any pair (x, y) with $y \neq x$ is not proximal. If all points are distal, we say that the action of G on X is distal or G is distal on X.

Remark 5.1 Our definition of distal as well proximal is suitable for action on compact spaces and algebraic actions. A more suitable definition would be using uniformity, [15].

Example 5.2 For $\Omega = \{1, 2, \ldots, p\}^{\mathbb{Z}}$ and S is the shift map $Sw(n) = w(n+1)$, we have for a $w \in \Omega$, any $w' \in \Omega$ for which $w' = w$ for arbitrarily large intervals is proximal. Thus, for $w \in \Omega$, there is a dense set of points that are proximal but certainly not all points.

Example 5.3 $x \longmapsto x + 1$ defines a proximal action of $\mathbb{Z}$ on $\mathbb{R} \cup \{\infty\}$ which is also the projective line.

Example 5.4 The $SL_2(\mathbb{R})$-action on the projective line is proximal and minimal.

Example 5.5 Isometries and equicontinuous actions are distal.

Example 5.6 Unipotent linear maps, orthogonal linear maps act distally.

Example 5.7 $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is distal on $\mathbb{T}^2$.

Exercise 5.8 (*Basic Properties*)

1. Let X_i be G_i spaces. Suppose the action of G_i on X_i are distal. Prove that the action of $\prod G_i$ on $\prod X_i$ is distal. Is the result true for proximal actions?
2. Let X and Y be G-spaces and $\phi : X \longrightarrow Y$ be a continuous surjective G-map—Y is known as *factor* of X. Suppose G is proximal on X. Prove that G is proximal on Y. Is the result true for distal actions?—here X and Y are general metric spaces.
3. Let X be a G-space and H be a subgroup of G. If G is distal on X, then H is distal on X. The converse is true, provided G/H is compact.
4. Let X be a G-space and H be a subgroup of G. Suppose G is proximal on X and G/H is compact. Then H is proximal on X.

Proposition 5.9 *A minimal subset for a proximal homeomorphism of a compact topological space is a fixed point.*

Proof Let $f : X \longrightarrow X$ be a homeomorphism of a compact space X. Suppose (X, f) is minimal and proximal. Let $x \in X$. Then, $d(f^{k_n}(x), f^{k_n+1}(x)) \to 0$ and $f^{k_n}(x) \to a$. This implies that $f(a) = a$. Since X is minimal $x = \{a\}$. □

Exercise 5.10 (*Proximality*)

1. Show that any proximal homeomorphism of a compact topological space has unique minimal set.

2. Determine all projective linear maps on the projective line that are proximal.

We now recall enveloping semigroup. Consider X^X, the space of all maps from X to X with product topology. If X is compact, then X^X is compact. If G acts on X, then G embeds inside X^X and the closure of G in X^X is called the enveloping semigroup, denoted by $E(X)$.

Theorem 5.11 (Ellis) *[6] The following are equivalent:*

1. G is distal on X.
2. $E(X)$ is a group.
3. For every $x \in X$, the orbit closure $\overline{Gx}$ is a closed G-invariant minimal set.

In this situation, $X = \cup E_i$, each E_i is a G-invariant minimal set and E_i's are disjoint.

We now consider the case where G acts on a group X by automorphisms. In this case, G is distal on X if and only if $e \notin \overline{Gx}$ for any $x \neq e \in X$. We have the following characterisations for linear actions.

Proposition 5.12 *Let $G \subset GL(V)$.*

1. $\alpha \in G$ is distal on V if and only if the eigenvalues of α are of absolute value one—due to Moore [17].
2. G is distal on V iff each $\alpha \in G$ is distal on V. In this situation, G is contained in a compact extension of an unipotent group—due to Conze and Guivarch [5].

The structure theorem for minimal distal actions shows that these action are built-up from one point space by isometric extensions of the following type: due to Furstenberg [11] with metrizability assumption which was later removed by Ellis [9].

Let (X, T) and (Y, T) be dynamical systems on compact spaces X and Y. Suppose $\phi : X \longrightarrow Y$ is a surjective map such that

1. $\phi(tx) = t\phi(x)$;
2. there is a continuous real valued function ρ defined on $X_\phi = \{(x_1, x_2) \in X \times X \mid \phi(x_1) = \phi(x_2)\}$ that is T-invariant;
3. ρ defines a metric on $X_y = \{x \mid \phi(x) = y\}$;
4. (X_y, ρ_y) are compact and are isometric.

In this case, we say that X is an isometric extension of Y.

Proposition 5.13 *The isometric extension of a distal (flow) system is distal.*

Proof Let $x_1, x_2 \in X$ such that closure of $\{(tx_1, tx_2) \mid t \in T\}$ meets the diagonal. Then the same is true under ϕ. Since T is distal on Y, $x_1, x_2 \in X_y$ for some $y \in Y$. Also, $(tx_1, tx_2) \in X_\phi$ for all $t \in T$. Since X_ϕ is closed, the closure of $\{(tx_1, tx_2) \mid t \in T\}$ is also in X_ϕ. Since the closure of $\{(tx_1, tx_2) \mid t \in T\}$ meets the diagonal, 0 is a limit point of $\rho(tx_1, tx_2)$ as ρ is continuous on X_ϕ. But since ρ is T-invariant, $\rho(x_1, x_2) = \rho(tx_1, tx_2)$, $\rho(x_1, x_2) = 0$. Thus, $x_1 = x_2$. □

Minimal distal actions are considered as the topological version of ergodic actions which is measure theoretic. However, when one considers without minimality particularly in the algebraic situation, they are hereditarily anti-thetic [21] (see [4] also for Seethoff's results in this direction). See [22] for recent developments in distal algebraic actions on locally compact groups and the references cited therein.

Exercise 5.14 (*Algebraic actions*)

1. Can one have proximal action by an automorphism on a compact group?
2. A topological group is called distal if G action on itself by conjugation is distal. Prove that compact extension of distal groups are distal.

6 Topological Transitivity and Mixing

We look into properties that are quite divergent from distality. We consider the properties of topological transitivity and mixing. The concept of topological transitivity was first defined by G.D. Birkhoff in 1920. This property is one of the oldest and foremost studied dynamical property. In this section, we will only consider cascades or semi-cascades (X, f) where (X, d) is a perfect, compact metric space. We identify a singleton with the point it contains.

Definition 6.1 For any two nonempty, open $U, V \subset X$ and $x \in X$, define the *return times*:

$$\begin{aligned} N(x, V) = &\quad \{n \in \mathbb{N} : f^n(x) \in U\} &&= \{n \in \mathbb{N} : x \in f^{-n}(U)\}; \\ N(U, V) = &\{n \in \mathbb{N} : f^n(U) \cap V \neq \varnothing\} &&= \{n \in \mathbb{N} : U \cap f^{-n}(V) \neq \varnothing\}; \\ N(U, x) = &\quad \{n \in \mathbb{N} : x \in f^n(U)\} &&= \{n \in \mathbb{N} : f^{-n}(x) \cap U \neq \varnothing\}. \end{aligned}$$

Definition 6.2 A cascade or semi-cascade (X, f) is said to be *topologically transitive* if for every pair of nonempty open sets U, V in X, there is a $n \in \mathbb{N}$ such that $f^n(U) \cap V \neq \varnothing$. Equivalently, $U \cap f^{-n}(V) \neq \varnothing$.

Roughly, topological transitivity can be described as the eventuality of the neighbourhood of every point to visit every region of the phase space at some time.

Definition 6.3 The cascade or semi-cascade (X, f) is said to be *point transitive* if there is an $x_0 \in X$ such that $\overline{\mathcal{O}(x_0)} = X$, i.e., X has a dense orbit.

All such points with dense orbits are called *transitive points* and the set of transitive points in X is denoted as $Trans(f)$. Both these definitions of transitivity are equivalent, in a wide class of spaces, including all perfect, compact metric spaces.

Theorem 6.4 *If X has no isolated point, then point transitivity implies the transitivity of (X, f). The converse holds if X is separable and of second category.*

The following equivalent conditions for transitivity of (X, f) can be taken as an exercise.

Exercise 6.5 1. f is topologically transitive.
2. for every pair of nonempty open sets U and V in X, there is a positive integer n such that $f^{-n}(U) \cap V \neq \varnothing$.
3. for every pair of nonempty open sets U and V in X, $N(U, V) \neq \varnothing$.
4. for every nonempty open set $U \subset X$, $\cup_{n=1}^{\infty} f^n(U)$ is dense in X.
5. for every nonempty open set $U \subset X$, $\cup_{n=1}^{\infty} f^{-n}(U)$ is dense in X.
6. if $E \subset X$ is closed and $f(E) \subset E$, then $E = X$ or E is nowhere dense in X.
7. if $U \subset X$ is open and $f^{-1}(U) \subset U$, then $U = \varnothing$ or U is dense in X.
8. There exists $x \in X$ such that the orbit $\mathcal{O}(x)$ is dense in X, i.e., the set $Trans(f)$ of transitive points is nonempty.
9. The set $Trans(f)$ of transitive points equals $\{x : \omega(x) = X\}$ and it is a dense G_δ subset of X.

Remark 6.6 All transitive equicontinuous cascades on compact metric spaces are minimal.

Example 6.7 Let $\mathbb{T}^1$ be the unit circle and $g : \mathbb{T}^1 \longrightarrow \mathbb{T}^1$ be the irrational rotation, defined by $g(\theta) = \theta + \alpha$, where α is a fixed irrational multiple of 2π. Then $(\mathbb{T}^1, g)$ is transitive.

We note that this cascade is minimal. In fact, every minimal cascade is transitive.

Example 6.8 Let $f : [0, 1] \longrightarrow [0, 1]$ be defined as $f(x) = 1 - |2x - 1|$. Then $([0, 1], f)$ is transitive. This f is called the *tent-map*. Here, for any nonempty open J in $[0, 1]$, there exists $n \in \mathbb{N}$ such that $f^n(J) = [0, 1]$.

Exercise 6.9 Prove that for the tent map on $[0, 1]$:

(a) an element x has finite orbit if and only if x is a rational number in $[0, 1]$.
(b) $Trans(f)$ equals the set of all irrational numbers in $[0, 1]$.

Definition 6.10 For the cascade or semi-cascade (X, f), *the backward orbit* of $x \in X$ is denoted as $\mathcal{O}^-(x)$ and defined as,

$$\mathcal{O}^-(x) = \{y \in X : f^n(y) = x \text{ for some } n \in \mathbb{N}\}.$$

The concept of transitivity deals with denseness of some forward orbit, while the concept of minimality implies that every orbit is dense. What would result if we want every backward orbit to be dense? We discuss a few basics on this and recommend our readers to refer to [2, 18] for details on this.

Definition 6.11 A cascade or semi-cascade (X, f) is called *strongly transitive* if $\mathcal{O}^-(x)$ is dense for every $x \in X$.

Remark 6.12 If (X, f) is strongly transitive, then (X, f) is topologically transitive. And if (X, f) is minimal then (X, f) is strongly transitive. It is not difficult to see that for cascades (X, f), strongly transitive is equivalent to minimal. And hence this property becomes distinct when we are considering semi-cascades.

Exercise 6.13 For a semi-cascade (X, f) the following are equivalent:

1. The system is strongly transitive.
2. For every nonempty, open set $U \subset X$ and every point $x \in X$, there exists $n \in \mathbb{N}$ such that $x \in f^n(U)$.
3. For every nonempty, open set $U \subset X$ and every point $x \in X$, the set $N(U, x)$ is nonempty.
4. For every nonempty, open set $U \subset X$ and every point $x \in X$, the set $N(U, x)$ is infinite.

Definition 6.14 A cascade or semi-cascade (X, f) is *locally eventually onto* if and only if for any nonempty, open $U \subset X$ there exists $N \in \mathbb{N}$ such that $f^N(U) = X$.

Recall the example (6.8) of the tent-map. This is an example of a locally eventually onto semi-cascade.

Exercise 6.15 For (X, f) the following are equivalent.

1. (X, f) is locally eventually onto.
2. For all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $f^{-N}(x) = \{y \in X : f^N(y) = x\}$ is ϵ-dense in X for every $x \in X$.
3. For all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $f^{-n}(x)$ is ϵ-dense in X for every $x \in X$ and every $n \geq N$.

Remark 6.16 Note that any finite product of locally eventually onto systems will be locally eventually onto.

But, an analogous statement of the above remark cannot be said about strongly transitive or transitive systems. One can just consider the irrational rotation on $\mathbb{T}^1$ as an example. So it becomes a natural question as to when can transitivity be preserved under products. This leads to the concept of mixing in topological dynamics, though this concept has been inspired by the same named property from another aspect of dynamics—Ergodic theory. We consider this aspect in another chapter.

Definition 6.17 A cascade or semi-cascade (X, f) is said to be *mixing* if for every pair V, W of nonempty open sets in X, there is a $N > 0$ such that $f^n(V) \cap W$ is nonempty for all $n \geq N$. (X, f) is called *weakly mixing* if the product system $(X \times X, f \times f)$ is transitive.

Remark 6.18 All locally eventually onto systems are mixing, all mixing systems are weakly mixing, and all weakly mixing systems are transitive.

Exercise 6.19 For (X, f), the following are equivalent.

1. (X, f) is weak mixing.
2. For nonempty, open sets U, V, W in X, there exists $N \in \mathbb{N}$ such that $f^{-N}(U) \cap V \neq \varnothing$ and $f^{-N}(U) \cap W \neq \varnothing$.
3. For nonempty, open sets U, V, W in X, there exists $N \in \mathbb{N}$ such that $f^{N}(U) \cap V \neq \varnothing$ and $f^{N}(U) \cap W \neq \varnothing$.
4. For every $N \in \mathbb{N}$ the product system $(X^N, f^{(N)})$ is topologically transitive.

Note that the statement in (4) above is a well-known consequence of the Furstenberg Intersection Lemma.

Lemma 6.20 (Furstenberg Intersection Lemma, [12]) *For a cascade or semi-cascade (X, f), assume that $N(U, V) \cap N(U, U) \neq \varnothing$ for every pair of nonempty, open $U, V \subset X$. Then for all nonempty, open $U_1, V_1, U_2, V_2 \subset X$, there exist nonempty, open $U_3, V_3 \subset X$ such that*

$$N(U_3, V_3) \quad \subset \quad N(U_1, V_1) \cap N(U_2, V_2).$$

Proof $N(U_1, V_1) \neq \varnothing$ implies there exists $n_1 \in \mathbb{N}$ such that $U_0 = U_1 \cap f^{-n_1}(V_1)$ is nonempty and open. $N(U_0, U_2) \neq \varnothing$ implies there exists $n_2 \in \mathbb{N}$ such that $U = U_1 \cap f^{-n_1}(V_1) \cap f^{-n_2}(U_2)$ is nonempty and open. Since f is transitive, $f^{-n_1-n_2}(V_2)$ is nonempty and open.

$$\begin{aligned}
&N(U, U) \cap N(U, f^{-n_1-n_2}(V_2)) \\
&\subset N(U_1, f^{-n_2}(U_2)) \cap N(f^{-n_1}(V_1), f^{-n_1}(f^{-n_2}(V_2))) \\
&= N(U_1, f^{-n_2}(U_2)) \cap N(f^{n_1}(f^{-n_1}(V_1)), f^{-n_2}(V_2)) \\
&\subset N(U_1, f^{-n_2}(U_2)) \cap N(V_1, f^{-n_2}(V_2)).
\end{aligned}$$

Fix $n_0 \in N(U_1, f^{-n_2}(U_2)) \cap N(V_1, f^{-n_2}(V_2))$. With $n = n_0 + n_2$, the sets $U_3 = U_1 \cap f^{-n}(U_2)$, $V_3 = V_1 \cap f^{-n}(V_2)$ are nonempty and open.

Let $k \in N(U_3, V_3)$. Then $f^{-k}(V_3) \cap U_3 \neq \varnothing$. That is $f^{-k}(V_1) \cap f^{-n-k}(V_2) \cap U_1 \cap f^{-n}(U_2) \neq \varnothing$. Hence $k \in N(U_1, V_1) \cap N(f^{-n}(U_2), f^{-n}(V_2)) = N(U_1, V_1) \cap N(U_2, V_2)$.

As before, $N(f^{-n}(U_2), f^{-n}(V_2)) \subset N(U_2, V_2)$ and so our assertion follows. □

Exercise 6.21 For (X, f), the following are equivalent.

1. (X, f) is mixing.
2. For every nonempty, open set $U \subset X$, and $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $f^{-n}(U)$ is ϵ-dense in X for all $n \geq N$.
3. For every nonempty, open set $U \subset X$, and $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $f^{n}(U)$ is ϵ-dense in X for all $n \geq N$.

Definition 6.22 (X, f) is called *strongly product transitive* if for every positive integer k, the product system $(X^k, f^{(k)})$ is strongly transitive.

Exercise 6.23 For (X, f), the following are equivalent.

1. (X, f) is strongly product transitive.
2. For $\epsilon > 0$ and every finite subset $F \subset X$, there exists $N \in \mathbb{N}$ such that $f^{-N}(x)$ is ϵ-dense in X for all $x \in F$.
3. For $\epsilon > 0$ and every finite subset $F \subset X$, there exist infinitely many $N \in \mathbb{N}$ such that $f^{-N}(x)$ is ϵ-dense in X for all $x \in F$.
4. The collection of subsets

$$\{N(U, x) : x \in X \text{ and } U \text{ is nonempty, open in } X\}$$

of $\mathbb{N}$, generates a filter of subsets of $\mathbb{N}$.

All these properties defined above are related.

$$\begin{array}{c} \text{Locally Eventually Onto} \begin{array}{l} \Longrightarrow \text{Mixing} \\ \Longrightarrow \text{Strongly Product Transitive} \end{array} \Longrightarrow \text{Weak Mixing} \Longrightarrow \text{Transitive} \\ \text{Transitive} \Longleftarrow \text{Strongly Transitive} \begin{array}{l} \Longleftarrow \text{Minimal} \\ \Longleftarrow \text{Locally Eventually Onto} \end{array} \end{array}$$

The reverse implications do not hold here.
One of the features of transitivity is recurrence.

Proposition 6.24 *Let I be an interval in $\mathbb{R}$. Then, every transitive map on I has a dense set of periodic points.*

Remark 6.25 We note that in general even for a locally eventually onto system there need not be any periodic point. This is illustrated in a beautiful example due to Elon Lindenstrauss, described in [2].

Though we can still have some form of recurrence for transitive systems.

Proposition 6.26 *For a transitive (X, f), $\Omega(f) = X$. The converse is not true in general.*

There are examples of cascades such that every element is non-wandering, but no subsystem, except the trivial, is transitive; for example, the reflection map $1 - x$ on $[0, 1]$.

Proposition 6.27 *If (X, f) is strongly transitive but not minimal, then the set of non-recurrent points is dense in X.*

Recall that a set $S \subset \mathbb{N}$ is *syndetic* if there is a constant $L > 0$ such that for every $n \in \mathbb{N}$ we have $[n, n + L] \cap S \neq \varnothing$. $S \in \mathbb{N}$ is called *thick*, if for every $n \in \mathbb{N}$, there exists $a_n \in S$ such that $\{a_n, a_n + 1, \ldots, a_n + n\} \in S$. And $S \in \mathbb{N}$ is called *cofinite*, if there exists $N \in \mathbb{N}$, such that $\{N, N + 1, N + 2, \ldots\} \subset S$.

Recurrence is usually given in terms of the return times. We describe the type of return time sets given by these various properties of transitivity and mixing for any two nonempty, open $U, V \subset X$ bringing out the distinction in these properties.

	$N(U, V)$	$N(x, V)$	$N(U, x)$
transitive	infinite	infinite for dense set of $x \in X$	can be empty for some $x \in X$
strongly transitive	infinite	infinite for dense set of $x \in X$	infinite for all $x \in X$
weakly mixing	thick	thick for dense set of $x \in X$	can be empty for some $x \in X$
mixing	cofinite	thick for dense set of $x \in X$	can be empty for some $x \in X$
strongly product transitive	thick	thick for all $x \in X$	thick for all $x \in X$
minimal	syndetic	syndetic for all $x \in X$	syndetic for all $x \in X$
locally eventually onto	cofinite	thick for all $x \in X$	cofonite for all $x \in X$

Mixing also has a strong inter-relation with proximality. We define the below property for a general flow, and mention an equivalent condition for weakly mixing cascades.

Definition 6.28 Let (X, G) be a flow. The regionally proximal relation $Q(X) \subset X \times X$ is defined as

$$Q(X) = \Big\{(x, y) : \text{ there exist } x_i \to x,\ y_i \to y,\ \text{and } g_i \in G \text{ such that } \lim g_i x_i = \lim g_i y_i\Big\}.$$

The relation $Q(X)$ turns out to be closed, invariant, and reflexive, but not necessarily transitive.

Remark 6.29 The flow (X, G) is equicontinuous if and only if $Q(X) = \triangle$. In fact, when $Q(X)$ is an equivalence relation then the flow $(X/Q(X), G)$ is an equicontinuous flow.

Theorem 6.30 *Let (X, f) be minimal. Then (X, f) is weakly mixing if and only if the regional proximal relation $Q(X) = X \times X$.*

7 Summary

We have just presented some basic concepts in *Topological Dynamics* here. For more details, we encourage the enthusiastic reader to refer to [1, 3, 12, 14–16, 23–25].

The sister branches of measurable dynamics (ergodic theory) and topological dynamics have their origin in Classical Mechanics, where we have a smooth transformation of a manifold, which also preserves a measure on this manifold. Both these theories have a parallel and inter-dependent growth, and both also have strong analogies in various aspects of recurrence observed. It is important to study both these concepts individually as one often influences the other.

The topological concept of transitivity is closely related to the measurable concept of ergodicity. Though it is difficult to say which definition influenced the other. The concepts of mixing and weak mixing had their origins first in the measurable sense and were borrowed for displaying a similar recurrence in the topological sense. The topological concept of minimality has no analogous property in the measurable sense. In the case of distality, the topological version came first and the theory of measurable distality was strongly influenced by the topological results. The measurable concept of uniquely ergodic has no counterpart in topological case. The topological concepts of strongly transitive and strongly product transitive are relatively new, and it would be a good research to determine their analogies in the measurable domain.

There are some systems for which topological and ergodic properties are related much closer. A classical example is the geodesic flow on surfaces of constant negative curvature.

We note that the best reference for the study of minimal systems is [3]. Much of the theory of enveloping semigroups was single-handedly developed by Ellis in [8]. We have just noted the basics here. We encourage our readers to refer to [3, 8, 10] for details. A more recent presentation can be seen in [19].

In his seminal paper on disjointness in topological dynamics in 1967, Furstenberg [12] started a systematic study of transitive dynamical systems. This laid a foundation for the classification of dynamical systems by their recurrent properties, which has been very useful in proving many combinatorial problems. We have described one such application in the proof of van der Waerden's theorem. But these recurrent properties also have applications in many Ramsey type results. The Auslander-Ellis Theorem and Furstenberg intersection lemma especially have been elementary tools in many problems concerning diophantine approximations. Readers can refer to [12] for some introductory details.

References

1. Akin, E. (1997). *Recurrence in topological dynamics: Furstenberg families and Ellis actions*. The University series in mathematicsNew York: Plenum Press.
2. Akin, E., Auslander, J., & Nagar, A. (2016). Variations on the concept of topological transitivity. *Studia Mathematica, 235*(3), 225–249.

3. Auslander, J. (1988). *Minimal flows and their extensions*, North-Holland mathematics studies (Vol. 153).
4. Brown, J. R. (1976). *Ergodic theory and topological dynamics*, Pure and applied mathematics (Vol. 70). New York: Academic Press [Harcourt Brace Jovanovich, Publishers].
5. Conze, J.-P., & Guivarc'h, Y. (1974). Remarques sur la distalité dans les espaces vectoriels. *Comptes rendus de l'Académie des Sciences Paris Séries A, 278*, 1083–1086.
6. Ellis, R. (1958). Distal transformation groups. *Pacific Journal of Mathematics, 8,* 401–405.
7. Ellis, R. (1960). A semigroup associated with a transformation group. *Transactions of the American Mathematical Society, 94*, 272–281.
8. Ellis, R. (1969). *Lectures on topological dynamics*. New York: W. A. Benjamin Inc.
9. Ellis, R. (1978). The Furstenberg structure theorem. *Pacific Journal of Mathematics, 76*, 345–349.
10. Ellis, D. B., & Ellis, R. (2014). *Automorphisms and equivalence relations in topological dynamics*, London mathematical society lecture note series (Vol. 412). Cambridge University Press.
11. Furstenberg, H. (1963). The structure of distal flows. *American Journal of Mathematics, 85*, 477–515.
12. Furstenberg, H. (1967). Disjointness in ergodic theory, minimal sets and a problem in diophantine approximation. *Mathematical Systems Theory, 1*, 1–49.
13. Furstenberg, H., & Weiss, B. (1978). Topological dynamics and combinatorial number theory. *Journal d'Analyse Mathematique, 34*, 61–85.
14. Glasner, E. (2003). *Ergodic theory via joinings*, Mathematical surveys and monographs (Vol. 101). AMS Providence.
15. Gottschalk, W. H., & Hedlund, G. A. (1955). *Topological dynamics* (Vol. 36). Providence: AMS American Mathematical Society Colloquium Publications.
16. Hasselblatt, B., & Katok, A. (2003). *A first course in dynamics: With a panorama of recent developments*. New York: Cambridge University Press.
17. Moore, C. C. (1968). Distal affine transformation groups. *American Journal of Mathematics, 90,* 733–751.
18. Nagar, A., & Kannan, V. (2003). *Topological transitivity for discrete dynamical systems*. Applicable mathematics in the golden age (pp. 534–584). Narosa Publications.
19. Nagar, A., & Singh, M. (2018). Topological dynamics of enveloping semigroups. arXiv:1810.12854.
20. Numakura, K. (1952). On bicompact semigroups. *Mathematical Journal of Okayama University, 1*, 99–108.
21. Raja, C. R. E. (2009). Distal actions and ergodic actions on compact groups. *New York Journal of Mathematics, 15*, 301–318.
22. Raja, C. R. E., & Shah, R. (2019). Some properties of distal actions on locally compact groups. *Ergodic Theory and Dynamical Systems, 39*, 1340–1360.
23. Vries de, J. (1993). *Elements of topological dynamics*, Mathematics and its applications (Vol. 257). Dordrecht: Kluwer.
24. Vries de, J. (2014). *Topological dynamical systems: An introduction to the dynamics of continuous mappings*, de Gruyter studies in mathematics (Vol. 59).
25. Walters, P. (1982). *An introduction to ergodic theory*. New York: Springer.

Basic Ergodic Theory

C. S. Aravinda and Vishesh S. Bhat

1 Introduction

These notes are based on the course of six lectures given by the first named author at the well-run workshop organised at IIT-Delhi in the month of December, 2017. The lectures were intended to be self-contained covering some basic facts in ergodic theory including a discussion of the Birkhoff ergodic theorem which, in a sense, heralded the beginning of ergodic theory. Since the audience mainly consisted of graduate students with different mathematical backgrounds, the lectures began with a quick recap of the construction of the Lebesgue measure in $\mathbb{R}$ and progressed gradually to a discussion of more general measures. After setting up the groundwork on measure preserving transformations and flows on measure spaces, the notion of ergodicity was introduced.

Following a brief look at a couple of illustrative examples of dynamical systems, the focus shifted to a discussion of one of the early interesting examples of an ergodic system, namely the geodesic flow on closed surfaces of constant negative curvature. This necessitated a working recapitulation of the geometry of the upper-half plane with respect to the hyperbolic metric, the lectures culminated with a sketch of a proof, due to Eberhard Hopf, of the ergodicity of the geodesic flow in this setting.

This author acknowledges the support of the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.01-0520.

C. S. Aravinda (✉)
TIFR - Centre for Applicable Mathematics (TIFR-CAM), Bengaluru, India
e-mail: aravinda@math.tifrbng.res.in

V. S. Bhat
Mechanics and Materials Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Japan
e-mail: vishesh.bhat@oist.jp

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_3

The notes, naturally, reflect the dynamics that the lectures carried and also include some historical titbits in an attempt to capture the significance of the exciting developments, that have shaped this field of study.

The first named author would like to record his deep gratitude to the organisers of this extremely well-run workshop, and to Nikita Agarwal who cheerfully conducted the afternoon tutorials at the workshop with great energy and lot of prior planning. Both the authors thank the efficient editors of this volume for their invitation to script the sketchy lecture notes into a coherent narrative, and the anonymous referees whose careful comments as well as suggestions to add a few explanatory lines at a couple of places helped weed out the several inadvertent typos and in improving the readability. The authors take full responsibility for any errors that may still remain despite their sincere efforts to make these notes error free.

2 Measure Theoretic Preliminaries

This section seeks to develop some rudimentary aspects of measure, starting with the illuminating case of the Lebesgue measure on the real line. Finding the measure of a set means to get a certain estimation of its size. A finite set could be measured by its cardinality, whereas what distinguishes an infinite set from a finite set is its intriguing property of being in bijective correspondence with a proper subset of itself. This begs the question as to how one would determine the size of an infinite subset of the real line $\mathbb{R}$?

For a subset which is an interval $I = (a, b) \subset \mathbb{R}$, its length $|I|$, namely $b - a$ seems a natural and a reasonable estimation of its size. In fact, the seminal investigation that Henri Lebesgue undertook culminating with the description of the so called Lebesgue measure, by exploiting the notion of length, appears in his fundamental paper of 1904 [10].

The basic idea of the Lebesgue measure on $\mathbb{R}$ stems from an effort to adapt the notion of length for an arbitrary subset of $\mathbb{R}$. This turns out to be a very profitable enterprise, as building on finer and subtle variants of this notion, allows one to describe a whole family of s-dimensional Hausdorff measures for each $s \in (0, 1]$; in turn giving rise to the notion of Hausdorff dimension of a given subset. We shall quickly uncover the main facets in this section, particularly mentioning the succinct and elegant work of Caratheodory [4].

We begin by first recalling the notion of outer measure.

Definition 2.1 If $A \subseteq \mathbb{R}$, the *(Lebesgue) outer measure* of A is

$$\mu^*(A) = \inf \left\{ \sum_{k=1}^{\infty} |I_k| \,:\, A \subseteq \bigcup_{k=1}^{\infty} I_k, \text{ where } (I_k)_{k=1}^{\infty} \text{ is a collection of open intervals} \right\}.$$

The completeness property of the reals ensures that if at least one of the members of the above set is finite, then μ^* will be a finite non-negative real number. If no such finite number exists, then the outer measure of A is said to be infinite.

Definition 2.2 If $A \subseteq \mathbb{R}$ and $h \in \mathbb{R}$, the *translate* of A by h is

$$A + h = \{x + h \,:\, x \in A\}.$$

The outer measure on $\mathbb{R}$ exhibits the following properties which can easily be derived from first principles.

Theorem 2.3 *This theorem features the basic properties of outer measure on* $\mathbb{R}$.

1. *(Non-negativity)* $0 \leq \mu^*(A) \leq +\infty$.
2. *(Monotonicity)* $A \subseteq B \Longrightarrow \mu^*(A) \leq \mu^*(B)$.
3. *(Countable subadditivity)* $A \subseteq \bigcup_{n=1}^{\infty} A_n \Longrightarrow \mu^*(A) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.
4. *(Translation invariance)* $\mu^*(A + h) = \mu^*(A)$.
5. $\mu^*(A) = |A|$, *the length of* A, *if* A *is an interval.*

While the above mentioned properties inherently follow from the definition; one other natural and desirable property is to expect that the outer measures of two disjoint sets A and B add up to the outer measure of their disjoint union $A \cup B$. This expectation lies at the heart of our discussion and, in a sense, the real essence of the theory lies in understanding this rather innocuous requirement.

A moment's reflection on what the finite additivity property ensures, can be gathered from the following. If $\{A_i\},\ i = 1, \ldots, \infty$ is a countable collection of pairwise disjoint subsets of $\mathbb{R}$, then

$$\sum_{i=1}^{n} \mu^*(A_i) = \mu^* \left(\bigcup_{i=1}^{n} A_i \right) \leq \mu^* \left(\bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \mu^*(A_i).$$

Taking limits as $n \to \infty$ on both sides results in the countable additivity of the outer measure.

But, the outer measure μ^* defined above has a singular shortcoming in that it is not finitely additive! One way to see this fact, a posteriori, is to glean from Vitali's construction in 1905 [12], of a non-measurable subset of $\mathbb{R}$. Recall that Vitali exhibited a proper non-empty subset C of $\mathbb{R}$, taking rational translates of which, one obtains a countable collection of pairwise disjoint subsets of $\mathbb{R}$. It is on this collection, that the outer measure μ^* cannot be countably additive. In particular, there are disjoint subsets A and B of $\mathbb{R}$ such that $\mu^*(A \cup B) \neq \mu^*(A) + \mu^*(B)$.

In other words, there are subsets X and O of $\mathbb{R}$ such that for the partition by O of X, into disjoint subsets $X \cap O$ and $X \cap O^c$, one has

$$\mu^*(X) \neq \mu^*(X \cap O) + \mu^*(X \cap O^c).$$

To see this, consider disjoint sets A and B and take $X = A \cup B$ and $O = A$. Therefore, $\mu^*(X) = \mu^*(A \cup B) \neq \mu^*(A) + \mu^*(B) = \mu^*(X \cap O) + \mu^*(X \cap O^c)$.

Consequently, one looks at the collection, $\mathfrak{M}$, of all those sets $E \subseteq \mathbb{R}$ such that

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c), \ \forall A \subseteq \mathbb{R}. \tag{1}$$

On this collection, $\mathfrak{M}$, the outer measure μ^* is countably additive. The collection $\mathfrak{M}$, which includes open intervals, constitutes a σ-algebra, and the outer measure restricted to $\mathfrak{M}$ is called the Lebesgue measure on $\mathfrak{M}$. The expression (1) is termed as the Caratheodory criterion and naturally leads to the definition of a (Lebesgue) measurable set. The next two definitions make this observation precise.

Definition 2.4 A family of subsets, $\mathfrak{M}$ of a set X is said to be a *σ-algebra* if the following hold:

1. $X \in \mathfrak{M}$;
2. $A \in \mathfrak{M} \implies A^c \in \mathfrak{M}$;
3. $\{A_i\}_{i=1}^{\infty} \in \mathfrak{M} \implies \bigcup_{i=1}^{\infty} A_i \in \mathfrak{M}$.

Definition 2.5 A set $E \subseteq \mathbb{R}$ is said to be *Lebesgue measurable* or *measurable* if the Caratheodory criterion (1) holds with respect to E.

In light of the preceding definitions, the conclusions of the next proposition can be deduced using properties of the outer measure given in Theorem 2.3.

Proposition 2.6

1. *If I is an interval, then $I \in \mathfrak{M}$ and $\mu^*(I) = |I|$.*
2. *If $A \in \mathfrak{M}$, then $A^c \in \mathfrak{M}$.*
3. *If $A, B \in \mathfrak{M}$, then $A \cup B,\ A \cap B \in \mathfrak{M}$.*
4. *If pairwise disjoint sets $A_1, A_2, \ldots, A_N \in \mathfrak{M}$ and $E \subseteq \mathbb{R}$, then*

$$\mu^*\left(E \cap \left(\bigcup_{k=1}^{N} A_k\right)\right) = \sum_{k=1}^{N} \mu^*(E \cap A_k).$$

5. *(Countable additivity or σ-additivity) If $\{A_n\}_{n=1}^{\infty}$ is any sequence of measurable sets, then $\bigcap_{n=1}^{\infty} A_n$ and $\bigcup_{n=1}^{\infty} A_n$ are also measurable. Further, if $\{A_n\}_{n=1}^{\infty}$ is a sequence of pairwise disjoint measurable sets, then $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$ and*

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

Definition 2.7 Suppose $A \in \mathfrak{M}$. Then its *(Lebesgue) measure*, $\mu(A)$ is defined to be its outer measure: $\mu(A) = \mu^*(A)$.

Remark 2.8

- The reason for the need of two different concepts is that both have their disadvantages.
- μ is an additive measure, but is not defined for all subsets of $\mathbb{R}$.
- μ^* is defined for all subsets of $\mathbb{R}$, but is not additive, as demonstrated by Vitali's construction.

A more restricted class of Lebesgue measurable sets are the Borel measurable sets.

Definition 2.9 If X is any topological space (in this case $\mathbb{R}$), then the σ-algebra, $\mathfrak{B}$ generated by the class of open sets in X (resp. open intervals in $\mathbb{R}$) are called the *Borel sets* of X (resp. $\mathbb{R}$).

Remark 2.10 It can be easily shown that the Borel σ-algebra for $\mathbb{R}$ includes the half-open intervals such as $[a, b)$ as well as closed intervals and further that every Borel set is (Lebesgue) measurable.

The important properties of the outer measure μ^* continue to hold on replacing μ^* by μ whenever $A \in \mathfrak{M}$.

Theorem 2.11 *Here, we enlist some additional properties of measurable sets.*

1. *Continuity: Suppose $A_1 \supseteq A_2 \supseteq A_3 \cdots$ and $B_1 \subseteq B_2 \subseteq B_3 \cdots$ are sequences of measurable sets, and $\mu(A_1) < \infty$. Then,*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) \quad \text{and} \quad \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n).$$

2. *Approximation: If $A \in \mathfrak{M}$, and $\mu(A) < \infty$, then for all $\epsilon > 0$ there exists a bounded closed set B and an open set C such that $B \subseteq A \subseteq C$, and $\mu(C \cap B^c) < \epsilon$.*

The previously sketched discussion of the construction of the Lebesgue measure on $\mathbb{R}$, starting from the notion of outer measure is, in a sense, a proto for the construction of measures more generally on complete metric spaces. In the setting of a metric space X together with the distance function d, one starts with the notion of a 'metric outer measure' which estimates the size of a subset A, by considering covers of A by a countable number of open balls; then, using radii of open balls, one considers an appropriate measure of their sizes to analogously replace lengths of intervals.

We shall elaborate more on this later when discussing Hausdorff measures, but will now proceed to a discussion of measures in general.

Definition 2.12 A *measure space* is a triple $(X, \mathfrak{M}, \mu)$, where X is any set, $\mathfrak{M}$ is a σ-algebra of measurable sets and μ is a σ-additive measure.

A measurable space is just the pair $(X, \mathfrak{M})$ with no specification about the measure. The concept of σ-finiteness is another desirable property for a measure to possess.

Definition 2.13 A measure space $(X, \mathfrak{M}, \mu)$ is said to be σ*-finite* if X can be written as a countable union of measurable sets of finite measure i.e., $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu(A_n) < +\infty$, for all n. μ is then said to be a σ*-finite measure*.

Definition 2.14 Given a measure space $(X, \mathfrak{M}, \mu)$, a set $A \subset X$ is said to be a *null set* or a set of measure zero if there exists a set $A_1 \in \mathfrak{M}$ so that $A \subseteq A_1$ and $\mu(A_1) = 0$. Furthermore, two sets $A_1, A_2 \subset X$ are said to be *equivalent mod* 0 if their symmetric difference, $A_1 \Delta A_2$ i.e., $(A_1 \setminus A_2) \cup (A_2 \setminus A_1)$ has measure zero and this is denoted as $A_1 \equiv A_2 \pmod 0$.

Remark 2.15

1. It should be noted in this context that not every measurable set is a Borel set. In fact, it is possible to construct sets of measure zero which are Lebesgue measurable but not Borel measurable. Thus, the Lebesgue measure serves as a completion of the Borel measure.
2. Note that a more formal definition of a complete measure is as follows: Given a measure space $(X, \mathfrak{M}, \mu)$, μ is complete if and only if for any $N \in \mathfrak{M}$ where $\mu(N) = 0$, $E \subseteq N$ implies $E \in \mathfrak{M}$. The Lebesgue measure is complete precisely in the above sense.

Another example of a finite measure space is the probability space which is the space of choice for ergodic theory. For a measure space $(X, \mathfrak{M}, \mu)$, if $\mu(X) = 1$, then X is a said to be probability space and μ a probability measure.

Measure zero sets are very useful in characterising properties in measure theory.

Definition 2.16 A property P of points of a set $A \subseteq X$ is said to hold *almost everywhere* (a.e.) if the set of points of A which do not satisfy P form a set of measure zero.

2.1 Measurable Functions and Transformations

We now move on to the notion of a measurable function which closely mirrors the topological definition of a continuous function. The first definition is formulated in the setting of general measure spaces.

Definition 2.17 (*Measurable functions or transformations*) If $(X, \mathfrak{M})$ and $(Y, \mathfrak{N})$ are two measurable spaces, then a map $f : X \longrightarrow Y$ is *measurable* if $f^{-1}(A)$ is measurable i.e., $f^{-1}(A) \in \mathfrak{M}$ for every $A \in \mathfrak{N}$. Further, if X and Y are topological spaces, then $f : X \longrightarrow Y$ is said to be *(Borel-) measurable* if it is measurable with respect to the Borel σ-algebras of X and Y.

Remark 2.18 The above definition implies that every continuous function is (Borel-) measurable.

In the sequel, we use the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ with the usual conventions. To keep things simple, in the remaining part of this section, we restrict ourselves to extended real-valued functions defined on $\mathbb{R}$ (equipped with the usual Lebesgue measure), unless otherwise explicitly stated, although the statements hold in the more general setting of complete measure spaces.

Remark 2.19 In particular, if $f : (\mathbb{R}, \mathfrak{L}) \longrightarrow (\bar{\mathbb{R}}, \mathfrak{B})$, where $\mathfrak{L}$ is the Lebesgue σ-algebra, and f is measurable as in the Definition 2.17, then f is said to be Lebesgue measurable.

For extended real-valued functions f, g, denote

$$(f \wedge g)(x) = \min\{f(x), g(x)\}, \quad (f \vee g)(x) = \max\{f(x), g(x)\}.$$

Proposition 2.20 *Measurable functions satisfy the following notable properties:*

1. *Suppose f, g are measurable functions and $c \in \mathbb{R}$, then cf, $f + g$, fg, $|f|$, $f \wedge g$, $f \vee g$ are measurable.*
2. *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions and $\lim_{n\to\infty} f_n(x) = f(x)$, then f is measurable.*
3. *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions. Let $g(x) = \inf\{f_n(x)\}$ and $h(x) = \sup\{f_n(x)\}$. Then g and h are measurable.*

Definition 2.21 The *indicator function* of a set $A \subseteq \mathbb{R}$ is the function

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Definition 2.22 A *simple function* is a function of the form

$$f = a_1\chi_{A_1} + \cdots + a_n\chi_{A_n} \quad \text{where } a_i \in \mathbb{R},\ A_i \in \mathfrak{M} \text{ and } \mu(A_i) < \infty.$$

Definition 2.23 The *integral* of a simple function $f = \sum_{i=1}^{n} a_i\chi_{A_i}$ is

$$\int f\,d\mu = \int_{\mathbb{R}} f\,d\mu = \sum_{i=1}^{n} a_i\mu(A_i).$$

Definition 2.24 (*Integral of nonnegative measurable functions*) If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a nonnegative measurable function, then its integral is

$$\int f\,d\mu = \sup\left\{\int g\,d\mu \ :\ g \text{ is a simple function such that } 0 \le g \le f\right\}.$$

Proposition 2.25 *If f, g are nonnegative measurable functions and $a > 0$, then*

$$\int af\,d\mu = a\int f\,d\mu, \quad \int (f+g)\,d\mu = \int f\,d\mu + \int g\,d\mu.$$

Moreover, if $f \le g$, then

$$\int f\,d\mu \le \int g\,d\mu.$$

This additivity property will allow us to extend the definition of integration to functions that change sign.

Definition 2.26 For an extended real-valued function f, define functions

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0, \\ 0 & \text{if } f(x) \le 0; \end{cases} \qquad f^-(x) = \begin{cases} -f(x) & \text{if } f(x) < 0, \\ 0 & \text{if } f(x) \ge 0. \end{cases}$$

Note that f^+ and f^- are nonnegative. They are measurable if f is, and $f = f^+ - f^-$, $|f| = f^+ + f^-$.

Definition 2.27 A measurable function is *integrable* if $\int |f|\,d\mu < +\infty$.

Definition 2.28 If f is an integrable function, its *integral* is

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu.$$

Definition 2.29 The *limit supremum* of a sequence is the least upper bound of the set of all subsequential limits of the sequence. That is,

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} (\sup\{a_m : m \ge n\}) = \inf_{n\ge 0}\left(\sup_{m\ge n} a_m\right).$$

Similarly, we define

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} (\inf\{a_m : m \ge n\}).$$

Theorem 2.30 (Fundamental convergence theorems) *Here, we record the fundamental convergence theorems in analysis, that we use in the sequel.*

1. *(Lebesgue's dominated convergence theorem) Suppose $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions and $\lim_{n\to\infty} f_n(x) = f(x)$, for all $x \in \mathbb{R}$, and $|f_n(x)| \le g(x)$ for all $n \in \mathbb{N}$, $x \in \mathbb{R}$ where g is an integrable function. Then,*

$$\lim_{n\to\infty}\int f_n\,d\mu = \int f\,d\mu.$$

2. *(Monotone Convergence Theorem) Suppose $(f_n)_{n=1}^{\infty}$ is a non-decreasing sequence of non-negative measurable functions $0 \le f_1 \le f_2 \le \cdots$. Let $f(x) = \lim_{n\to\infty} f_n(x)$.*

Then,

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

3. *(Fatou's Lemma) If* $(f_n)_{n=1}^{\infty}$ *is a sequence of nonnegative measurable functions, then*

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu.$$

Definition 2.31 Two functions f and g are said to be equal *almost everywhere*, written $f = g$ a.e., if $\{x \, : \, f(x) \neq g(x)\}$ is a set of measure zero.

Proposition 2.32 *If* f *is a function on a Lebesgue measurable set* E *and* $g = f$ *a.e., then* g *is Lebesgue measurable if and only if* f *is Lebesgue measurable.*

Definition 2.33 Consider the set of all integrable functions on $\mathbb{R}$. The function space L^1 is the set of all equivalence classes of integrable functions on $\mathbb{R}$, where we set $f \simeq g$ if $f = g$ a.e. The L^1 *norm* is given by

$$\|f\|_1 := \int |f| \, d\mu.$$

Theorem 2.34 L^1 *is complete, i.e., given a Cauchy sequence* $\{f_n\}_{n=1}^{\infty}$ *in* L^1*, there exists* $f \in L^1$ *such that* $\lim_{n\to\infty} \|f_n - f\|_1 = 0$.

Generalising the L^1 notion to functions on arbitrary complete measure spaces, we have the following definition.

Definition 2.35 Let $(X, \mathfrak{M}, \mu)$ be a complete measure space and $f : X \longrightarrow \bar{\mathbb{R}}$ be a measurable function, then for each integer $p \geq 1$, we say that $f \in L^p(\mu)$ if

$$\int_X |f|^p \, d\mu < \infty.$$

For any such $f \in L^p(\mu)$, we may define the L^p*-norm* as

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

Identifying the functions whose values agree a.e. allows for defining a metric on the space $L^p(\mu)$ by means of the L^p-norm. We treat $L^p(\mu)$ as the set of equivalence classes of functions which coincide a.e.. Thus, $L^p(\mu)$ becomes a Banach space for $1 \leq p < \infty$. In particular, $L^2(\mu)$ is a Hilbert space with the inner product defined by

$$\langle f, g\rangle := \int_X fg\, d\mu.$$

Definition 2.36 $f : X \longrightarrow \mathbb{R}$ is said to be *compactly supported* if the closure of the set of points in X where the value of f is non-zero, is a compact subset of X.

Notation 2.37 We denote the set of all compactly supported (real-valued) continuous functions on X as $C_c(X)$.

Theorem 2.38 (Lusin's Theorem) *If X is a locally compact Hausdorff topological space and if $f : X \longrightarrow \bar{\mathbb{R}}$ is a measurable function such that $f(x) = 0$, for all $x \notin A \subset X$, where $\mu(A) < \infty$, then given $\epsilon > 0$, there exists a $g \in C_c(X)$ so that*

$$\mu\left(\{x \ : \ f(x) \neq g(x)\}\right) < \epsilon.$$

Theorem 2.39 *For $1 \leq p < \infty$, $C_c(X)$ is dense in $L^p(\mu)$.*

Definition 2.40 Let $(X, \mathfrak{M})$ be a measurable space and $\mu,\ \nu : X \longrightarrow [0, \infty)$ be two measures on $\mathfrak{M}$. We say that μ is *absolutely continuous* with respect to ν if $A \in \mathfrak{M}$ and $\nu(A) = 0$ implies $\mu(A) = 0$. This is denoted as $\mu \ll \nu$.

Theorem 2.41 (Radon–Nikodym) *If $(X, \mathfrak{M}, \nu)$ is a σ-finite measure space, then $\mu \ll \nu$ if and only if there exists a function $f \in L^1(\nu)$ such that*

$$\mu(A) = \int_A f\, d\nu \ \text{ for every } A \in \mathfrak{M}.$$

The function f is unique a.e. with respect to ν and is written as $\frac{d\mu}{d\nu}$, called the Radon-Nikodym derivative *of μ w.r.t ν.*

2.2 Hausdorff Measures

In this subsection, we outline the notion of more general measures called Hausdorff measures that subsume the Lebesgue measure. It is assumed that (X, d) is a non-empty metric space. The notion of Hausdorff dimension of a subset $A \subset X$ arises from the construction of Hausdorff measures [6].

Definition 2.42 A function μ defined on $\mathcal{P}(X)$ is called a *metric outer measure* if it satisfies the following:

1. $\mu^*(A) \geq 0$, for all $A \in \mathcal{P}(X)$;
2. $\mu^*(\varnothing) = 0$;
3. (Monotonicity) $A_1 \subseteq A_2 \implies \mu^*(A_1) \leq \mu^*(A_2)$;

4. (Countable subadditivity) if $\{A_n\}_{n=1}^{\infty}$ is a countable collection of members of $\mathcal{P}(X)$, then $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$;
5. if $A_1, A_2 \in \mathcal{P}(X)$ with $d(A_1, A_2) > 0$, then $\mu^*(A_1 \cup A_2) = \mu^*(A_1) + \mu^*(A_2)$.

A familiar example of such an outer measure is the Lebesgue outer measure discussed in the earlier sections. Before defining the Hausdorff measure, we remark that as in the case of $\mathbb{R}$, a subset E of a space X is said to be measurable if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c), \quad \forall A \in \mathcal{P}(X).$$

The class of measurable sets in X evidently form a σ-algebra $\mathfrak{P}$ so that μ when restricted to $\mathfrak{P}$, is countably additive and thus a measure in the usual sense. We henceforth use the usual μ notation for the measure.

Definition 2.43 Given a metric space (X, d) and $A \subset X$, the diameter of A is given as $\delta(A) := \sup\{d(x, y) \ : \ x, y \in A\}$.

Let (X, d) be a metric space and let $\alpha(> 0) \in \mathbb{R}$. Let $A \subset X$. Given $\epsilon > 0$, consider

$$H_\alpha^\epsilon(A) = \inf\left\{\sum_{k=1}^{\infty} \delta(A_k)^\alpha \ : \ A \subseteq \bigcup_{k=1}^{\infty} A_k \text{ where } \delta(A_k) < \epsilon \ \forall k\right\},$$

the infimum being taken over all countable covers of the set A whose members have diameter less than ϵ. Note that if $\epsilon_1 < \epsilon$, then $H_\alpha^{\epsilon_1}(A) \geq H_\alpha^\epsilon(A)$. Therefore, $\lim_{\epsilon \to 0} H_\alpha^\epsilon(A)$ exists, though it may be infinite, and we write $H_\alpha(A) = \lim_{\epsilon \to 0} H_\alpha^\epsilon(A)$.

Theorem 2.44 *For each $\alpha > 0$, H_α is a metric outer measure on X called the Hausdorff outer measure of dimension α and when restricted to the σ-algebra of measurable sets, is called the Hausdorff measure of dimension α on X.*

Note that if $\alpha = 0$, then H_α is merely, the counting measure.

Theorem 2.45 *(i) If $H_\alpha(A) < \infty$, then $H_\beta(A) = 0$ for $\beta > \alpha$.*
(ii) If $H_\alpha(A) > 0$, then $H_\beta(A) = \infty$ for $\beta < \alpha$.

Proof It is easy to see that *(i)* and *(ii)* are equivalent. Therefore, we prove *(i)*. Suppose $A = \bigcup_{k=1}^{\infty} A_k$, with $\delta(A_k) < \epsilon$. If $\beta > \alpha$, then

$$H_\beta^\epsilon(A) \leq \sum_{k=1}^{\infty} \delta(A_k)^\beta \leq \epsilon^{\beta-\alpha} \sum_{k=1}^{\infty} \delta(A_k)^\alpha.$$

That is, $H_\beta^\epsilon(A) \leq \epsilon^{\beta-\alpha} H_\alpha^\epsilon(A)$. Letting $\epsilon \to 0$, we see that $H_\beta(A) = 0$ if $H_\alpha(A) < \infty$. □

As a consequence of the above theorem, for $A \subset X$, there exists $d \in \mathbb{R}$ such that

$$\begin{cases} H_m(A) = 0 & \text{if } m > d, \\ H_m(A) = \infty & \text{if } m < d. \end{cases}$$

The d, obtained as above, is called the *Hausdorff dimension* of the set A, denoted by $\mathcal{H}_{dim}(A)$.

Example 2.46

1. If A is any countable set then, $\mathcal{H}_{dim}(A) = 0$.
2. If $X = \mathbb{R}$ and $\alpha = 1$, then it is straightforward to check that H_1 is the Lebesgue measure.
3. The Cantor ternary set is an example of an uncountable set of zero Lebesgue measure, as opposed to countable sets which are also of Lebesgue measure zero. It can be shown that its Hausdorff dimension is $\dfrac{\ln 2}{\ln 3}$.

If $X = \mathbb{R}^n,\ n > 1$, then H_n is not the same as the Lebesgue measure, but is *comparable* to it, a fact elucidated in the next theorem.

Theorem 2.47 *Let $A \subset \mathbb{R}^n$.*

1. *Then, there exists positive constants C_1 and C_2 depending only on the dimension n such that*
$$C_1 H_n(A) \leq \lambda(A) \leq C_2 H_n(A),$$
for $A \subset \mathbb{R}^n$, λ being the Lebesgue measure on $\mathbb{R}^n$.
2. *If $\alpha > n$, then $H_\alpha(A) = 0$, for every $A \subset \mathbb{R}^n$.*

3 Recurrence and Ergodic Theorems

Let $(X, \mathfrak{M}, \mu)$ be a measure space. A transformation $T : X \longrightarrow X$ is said to be a *measurable transformation* (with respect to μ) if the inverse image of every μ-measurable set is μ-measurable. And a μ-measurable transformation T of X into itself is said to be *measure preserving* if $\mu(T^{-1}(E)) = \mu(E)$ for every μ-measurable subset E of X.

Example 3.1

1. Let $X = [0, 1)$ and λ be the Lebesgue measure on X. Let $c \in X$ be any point. Then the transformation $T : X \longrightarrow X$ defined by $T(x) = x + c \pmod 1$ is measure preserving.

2. Let $X = [0, 1)$ and λ be the Lebesgue measure on X. Define $T : X \longrightarrow X$ as

$$T(x) = \begin{cases} 2x & \text{for } 0 \leq x < \frac{1}{2} \\ 2x - 1 & \text{for } \frac{1}{2} \leq x < 1. \end{cases}$$

   It can be easily verified that T as defined above is a measure preserving transformation.
3. Given $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ where $\mathbb{R}^n$ is equipped with the usual Lebesgue measure. The affine transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined as $T(x) = x + a$ is invertible and measure preserving.

In the context of ergodic theory, a measurable space $(X, \mathfrak{M}, \mu)$ equipped with a measure preserving transformation T constitutes a *dynamical system* denoted by $(X, \mathfrak{M}, \mu, T)$.

3.1 Recurrence

In the sequel, we assume that $(X, \mathfrak{M}, \mu)$ is a probability space i.e. $\mu(X) = 1$. Given a measure preserving transformation T on a measure space $(X, \mathfrak{M}, \mu)$, T is said to be recurrent if for any given set of positive measure $A \subset X$, almost all points of A return to A after at most finitely many iterations of T.

Theorem 3.2 (Poincare recurrence theorem) *Let $(X, \mathfrak{M}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure preserving transformation. Given $A \in \mathfrak{M}$, let A_0 be the set of points $x \in A$ such that $T^n(x) \in A$ for infinitely many $n \geq 0$. Then $A_0 \in \mathfrak{M}$ and $\mu(A_0) = \mu(A)$.*

Proof Let

$$C_n = \left\{ x \in A \ : \ T^k(x) \notin A \ \forall k \geq n \right\}.$$

Therefore $A_0 = A \setminus \bigcup_{n=1}^{\infty} C_n$. In order to prove the theorem, it is enough to show that

1. $C_n \in \mathfrak{M}$ and
2. $\mu(C_n) = 0$ for every $n \geq 1$.

1. Now, $C_n = A \setminus \bigcup_{k \geq n} T^{-k}(A)$. Since $T^{-k}(A) \in \mathfrak{M}$ for every $k \geq 1$, we see that $C_n \in \mathfrak{M}$.
2. Also,

$$C_n \subset \bigcup_{k \geq 0} T^{-k}(A) \setminus \bigcup_{k \geq n} T^{-k}(A)$$

$$\implies \mu(C_n) \leq \mu\left(\bigcup_{k \geq 0} T^{-k}(A) \right) - \mu\left(\bigcup_{k \geq n} T^{-k}(A) \right).$$

Now, observe that $\bigcup_{k\geq n} T^{-k}(A) = T^{-n}\left(\bigcup_{k\geq 0} T^{-k}(A)\right)$. Since T is measure preserving, this implies

$$\mu\left(\bigcup_{k\geq 0} T^{-k}(A)\right) = \mu\left(\bigcup_{k\geq n} T^{-k}(A)\right).$$

Therefore $\mu(C_n) = 0$. □

3.2 Birkhoff Ergodic Theorem and the Notion of Ergodicity

Let $(X, \mathfrak{M}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure preserving transformation. Let $E \in \mathfrak{M}$. Given $x \in X$, one would like to ask with what frequency do the elements of the set $\{x, Tx, T^2x, \ldots\}$ lie in the set E?

Clearly $T^i x \in E$ if and only if $\chi_E(T^i x) = 1$; therefore the number of elements of $\{x, Tx, T^2x, \ldots, T^{n-1}x\}$ in E is $\sum_{k=0}^{n-1} \chi_E(T^k x)$ or the relative number of $\{x, Tx, \ldots, T^{n-1}x\}$ in E is $\frac{1}{n}\sum_{k=0}^{n-1} \chi_E(T^k x)$.

Around the turn of the century, the work of Boltzmann and Gibbs on statistical mechanics raised a mathematical problem which can be stated as follows: Given a measure preserving transformation T of a probability space and an integrable function $f : X \longrightarrow \mathbb{R}$, find conditions under which

$$\lim_{n\to\infty} \frac{f(x) + f(Tx) + \cdots + f(T^{n-1}x)}{n}$$

exists and is constant almost everywhere.

In 1931 [3], Birkhoff proved that for any T and f, the above limit exists almost everywhere. From this, he concluded that a necessary and sufficient condition for its value to be constant almost everywhere, is that there exist no set $A \in \mathfrak{M}$ such that $0 < \mu(A) < 1$ and $T^{-1}A = A$. As we will see later, the fact that this limit is constant easily implies that it is equal to the integral of f over X. Transformations T which satisfy this condition are called *ergodic* and ergodic theory is essentially the study of such transformations. The Birkhoff Ergodic theorem is the first fundamental result that sets the tone for much of what follows.

Theorem 3.3 (Birkhoff Ergodic Theorem) *Let $(X, \mathfrak{M}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure preserving transformation. If $f \in L^1(\mu)$ then the limit*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x)) = \widetilde{f}(x),$$

exists for almost every point $x \in X$, $\widetilde{f} \in L^1(\mu)$ *and* $\widetilde{f} \circ T = \widetilde{f}$ *almost everywhere. Furthermore,*

$$\int_X \widetilde{f}\, d\mu \;=\; \int_X f\, d\mu.$$

If f is any measurable function, let $g(x) = f(Tx)$. Since T is measurable, the function g is measurable so that, writing $g(x) = Uf(x)$, the transformation U assigns to each measurable function f, a measurable function g. Clearly, U is linear and g is non-negative if f is so. Moreover, we have:

Theorem 3.4 *If* $1 \leq p \leq \infty$ *and* $\|f\|_p$ *denotes the* L^p*-norm of* f*, then* $\|g\|_p = \|f\|_p$ *for* $g = Uf$.

Proof Let $E \in \mathfrak{M}$ and $f = \chi_E$. Then $g = Uf = f(Tx) = \chi_{T^{-1}(E)}$. Therefore,

$$\|g\|_p^p \;=\; \mu(T^{-1}(E)) \;=\; \mu(E) \;=\; \|f\|_p^p.$$

It follows that $\|g\|_p = \|f\|_p$ for every non-negative simple function. If f is any non-negative measurable function, there exists a sequence of simple, non-negative measurable functions $\{s_n\}_{n=1}^{\infty}$ such that $s_n \to f$, as $n \to \infty$, with $s_1 \leq s_2 \leq \cdots \leq f$. Now, since $t_n = Us_n$ is also an increasing sequence of simple functions, converging to g, monotone convergence theorem implies that

$$\|g\|_p \;=\; \lim_{n\to\infty} \|t_n\|_p \;=\; \lim_{n\to\infty} \|s_n\|_p \;=\; \|f\|_p.$$

The general case of f now follows by writing $f = f^+ - f^-$ and applying the above conclusion to f^+ and f^- separately. □

In particular, if $f \in L^2(\mu)$ we have showed that $g(x) = Uf(x) = f(Tx)$ is also in $L^2(\mu)$ and that $\|g\|_2 = \|f\|_2$. In other words, U is an isometric transformation of the Hilbert space $L^2(\mu)$ into itself.

If, in addition, T is invertible (i.e., there exists a measure preserving transformation $S : X \longrightarrow X$ such that $ST = TS = Id_X$) and if V is the isometric transformation in $L^2(\mu)$ corresponding to its inverse, then $UV = VU$ is the identity transformation in $L^2(\mu)$. Therefore, the range of V is the whole of $L^2(\mu)$; in other words, U is a unitary transformation in $L^2(\mu)$ and V is its inverse. Thus, an invertible measure preserving transformation on a measure space (X, μ) induces an invertible unitary transformation in the Hilbert space $L^2(\mu)$.

Therefore, in so far as it concerns functions $f \in L^2(\mu)$, the existence of the limit of the averages is reduced to the problem of existence of the limit as $n \to \infty$ of $\frac{1}{n}\sum_{k=0}^{n-1} U^k f(x)$, where U is an isometric transformation in the Hilbert space $L^2(\mu)$. Precisely, this convergence, known as the mean ergodic theorem, was proven by J. von Neumann in 1932 [13].

Theorem 3.5 (Mean ergodic theorem) *If U is an isometric transformation in an arbitrary Hilbert space H and if P is the orthogonal projection on the closed linear subspace of all $f \in H$ satisfying $Uf = f$, then $\frac{1}{n}\sum_{k=0}^{n-1} U^k f$ converges in norm as $n \to \infty$ to Pf for all $f \in H$.*

We will skip a proof of this and prove the more general Birkhoff ergodic theorem (BET, for short). We prove the first part of the BET and prove the more general L^p version of the second part as a corollary. The key step in the proof of BET is itself a useful lemma known as the Maximal ergodic theorem.

Lemma 3.6 (Maximal ergodic theorem) *Given $f \in L^1(\mu)$, put*

$$E(f) \;=\; \left\{ x \;:\; \max_{n\geq 0} \left(\sum_{k=0}^{n-1} f(T^k x) \right) > 0 \right\}.$$

Then $\int_{E(f)} f \, d\mu \geq 0$.

Proof Define

$$\begin{aligned} f_0 &:= 0, \\ f_n &:= f + f \circ T + f \circ T^2 + \cdots + f \circ T^{n-1} \\ &= f + Uf + U^2 f + \cdots + U^{n-1} f. \end{aligned}$$

Let $F_n = \max_{0\leq k\leq n} f_k$. Therefore

$$E(f) = \bigcup_{n-1}^{\infty} \{x \;:\; F_n(x) > 0\} = \bigcup_{n-1}^{\infty} E_n.$$

Clearly, $F_n \in L^1(\mu)$ and, for $0 \leq k \leq n$, we have $F_n \geq f_k$. Therefore $UF_n \geq Uf_k$ because $U : L^1(\mu) \longrightarrow L^1(\mu)$ is a positive linear operator (i.e., $f \geq 0$ implies $Uf \geq 0$) and hence,

$$UF_n + f \geq Uf_k + f = f_{k+1}.$$

In other words,

$$UF_n + f \geq \max_{1\leq k\leq n} f_k(x) \;=\; \max_{0\leq k\leq n} f_k(x) \;=\; F_n(x) \text{ when } F_n(x) \;>\; 0.$$

That is, $f \geq F_n - UF_n$ on $\{x \;:\; F_n(x) > 0\} = E_n$. Therefore,

$$\begin{aligned}\int_{E_n} f\,d\mu &\geq \int_{E_n} F_n\,d\mu - \int_{E_n} UF_n\,d\mu \\ &= \int_X F_n\,d\mu - \int_{E_n} UF_n\,d\mu \\ &\geq \int_X F_n\,d\mu - \int_X UF_n\,d\mu \\ &= 0.\end{aligned}$$

The second equality above holds because $F_n = 0$ on $X \setminus E_n$, the third inequality holds because $F_n \geq 0$ implies $UF_n \geq 0$ and the last equality holds because $\|U\| = 1$. Finally, since $E_1 \subseteq E_2 \subseteq \cdots$, we have that $E_n \to E(f)$ and we are done. □

Corollary 3.7 *If $A \subset E(f),\ A \in \mathfrak{M}$ and $T^{-1}A = A$, then,*

$$\int_A f\,d\mu \geq 0.$$

Proof Since $T^{-1}A = A$, we see that $E(f\chi_A) = A$. Therefore, the lemma above implies $0 \leq \int_{E(f\chi_A)} f\chi_A\,d\mu = \int_A f\chi_A\,d\mu = \int_A f\,d\mu$. □

Theorem 3.8 *Let $(X, \mathfrak{M}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure preserving transformation. If $f \in L^1(\mu)$, then the limit*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$$

exists for almost every point $x \in X$.

Proof For each $\alpha,\ \beta \in \mathbb{R}$ with $\alpha < \beta$, let

$$E_{\alpha,\beta} \;=\; \left\{ x \in X \;:\; \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) < \alpha < \beta < \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \right\}.$$

Clearly, $E_{\alpha,\beta} \in \mathfrak{M}$. We will show that $\mu(E_{\alpha,\beta}) = 0$ for each $\alpha,\ \beta$. This would imply that $\bigcup E_{\alpha,\beta}$, where $\alpha,\ \beta \in \mathbb{R}$ such that $\alpha < \beta$, has measure zero and hence the limit exists almost everywhere.

Put $f^*(x) = \sup_{n\geq 1} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ and $f_*(x) = \inf_{n\geq 1} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$. Therefore,

$$E_{\alpha,\beta} \subset \{x \;:\; f^*(x) > \beta\} = \{x \;:\; (f^* - \beta)(x) > 0\} = E(f - \beta)$$

and $E_{\alpha,\beta} \subset \{x : f_*(x) < \alpha\}$.

We first show that $E_{\alpha,\beta}$ is T-invariant. That is, we show that $T^{-1}(E_{\alpha,\beta}) = E_{\alpha,\beta}$.

Let $a_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$. Then, $\frac{n+1}{n}a_{n+1}(x) - a_n(Tx) = \frac{f(x)}{n}$. Therefore,

$$\limsup_{n\to\infty}(a_{n+1}(x) + \frac{1}{n}a_{n+1}(x) - a_n(Tx)) = \limsup_{n\to\infty}\frac{f(x)}{n}.$$

This implies that $\limsup_{n\to\infty}(a_{n+1}(x) - a_n(Tx)) = 0$. That is, $\limsup_{n\to\infty}(a_{n+1}(x)) = \limsup_{n\to\infty}(a_n(Tx))$. Similarly, $\liminf_{n\to\infty}(a_{n+1}(x)) = \liminf_{n\to\infty}(a_n(Tx))$.

Therefore, $T^{-1}(E_{\alpha,\beta}) = E_{\alpha,\beta}$.

By Corollary 3.7, we get $\int_{E_{\alpha,\beta}}(f-\beta)\,d\mu \geq 0$ or $\int_{E_{\alpha,\beta}} f\,d\mu \geq \beta\mu(E_{\alpha,\beta})$. Now $E_{\alpha,\beta} \subset \{x : f_*(x) < \alpha\} = \{x : -f_* > -\alpha\} = \{x : (-f)^* > -\alpha\}$.

Therefore, by the maximal ergodic theorem 3.6, $\int_{E_{\alpha,\beta}}(-f)\,d\mu \geq -\alpha\mu(E_{\alpha,\beta})$ or $\int_{E_{\alpha,\beta}} f\,d\mu \leq \alpha\mu(E_{\alpha,\beta})$. Thus, $\beta\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} f\,d\mu \leq \alpha\mu(E_{\alpha,\beta})$.

But $\alpha < \beta$. Therefore, the above inequality holds only if $\mu(E_{\alpha,\beta}) = 0$. □

Corollary 3.9 *(i) If $f \in L^p(\mu)$, $1 \leq p \leq \infty$, the function $\widetilde{f}$ defined by,*

$$\widetilde{f}(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$$

is in $L^p(\mu)$ and satisfies

$$\lim_{n\to\infty}\left\|\widetilde{f} - \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right\|_p = 0.$$

(ii) $\widetilde{f}(Tx) = \widetilde{f}(x)$.
(iii) For $f \in L^p(\mu)$, $\int_X \widetilde{f}\,d\mu = \int_X f\,d\mu$.

Proof *(i)* Since X is a probability space, $\mu(X) = 1$. Therefore, $f \in L^1(\mu)$ and $\widetilde{f}(x)$ makes sense. Moreover, $|f| \in L^1(\mu)$ and $\left|\widetilde{f}(x)\right| \leq \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left|f(T^k x)\right|$ for a.e. x (this limit exists since $|f| \in L^1(\mu)$). That is, $\left|\widetilde{f}(x)\right|^p \leq \lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=0}^{n-1}\left|f(T^k x)\right|\right)^p$. Since $\left|\widetilde{f}\right|^p \geq 0$,

$$\begin{aligned}\|\widetilde{f}\|_p^p = \int_X |\widetilde{f}|^p \, d\mu &= \int_X \lim_{n\to\infty} \left(\frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)|\right)^p d\mu \\ &= \int_X \liminf_{n\to\infty} \left(\frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)|\right)^p d\mu \\ &\le \liminf_{n\to\infty} \int_X \left(\frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)|\right)^p d\mu. \text{ (Fatou's Lemma)}\end{aligned}$$

Now

$$\begin{aligned}\int_X \left(\frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)|\right)^p d\mu &= \left\| \frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)| \right\|_p^p \\ &\le \left(\frac{1}{n}\sum_{k=0}^{n-1} \|f(T^k x)\|_p\right)^p \\ &= \left(\frac{1}{n}\sum_{k=0}^{n-1} \|f\|_p\right)^p \ (T^k \text{ is measure preserving}) \\ &= \|f\|_p^p.\end{aligned}$$

Therefore

$$\|\widetilde{f}\|_p^p \le \liminf_{n\to\infty} \int_X \left(\frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)|\right)^p d\mu \le \liminf_{n\to\infty} \|f\|_p^p = \|f\|_p^p < \infty,$$

since $f \in L^p(\mu)$. Therefore $\widetilde{f} \in L^p(\mu)$. □

Definition 3.10 (*Convergence in the L^p-norm*) Consider the case $f \in L^\infty(\mu)$, i.e., $\sup_{x\in X} |f(x)| < \infty$ a.e. Clearly, $f \in L^1(\mu)$ and the sequence of functions

$$\left| \widetilde{f} - \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) \right|^p$$

converges to 0 a.e. Moreover,

$$|\widetilde{f}(x)| \le \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |f(T^k x)| \le \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \|f\|_\infty = \|f\|_\infty \text{ -a.e.}$$

Therefore,

$$\left|\widetilde{f}(x)-\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)\right|^p \le \left|\|f\|_\infty+\frac{1}{n}\sum_{k=0}^{n-1}\left\|f\circ T^k\right\|_\infty\right|^p \le \left(2\|f\|_\infty\right)^p = \text{constant}.$$

Hence by dominated Convergence theorem,

$$\int_X\left|\widetilde{f}-\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)\right|^p d\mu\to 0 \text{ -a.e.}$$

That is, for $f\in L^p(\mu)$, $\lim_{n\to\infty}\left\|\widetilde{f}-\frac{1}{n}\sum_{k=0}^{n-1}f\circ T^k\right\|_p=0$. Now, let $f\in L^p(\mu)$ and let $\varepsilon>0$. There is an $f_0\in L^\infty(\mu)$ such that $\|f-f_0\|_p\le\varepsilon/3$ and there exists an $N>0$ such that $\left\|\widetilde{f_0}-\frac{1}{n}\sum_{k=0}^{n-1}f_0\circ T^k\right\|_p\le\varepsilon/3$ for $n\ge N$.

Then,

$$\begin{aligned}&\left\|\widetilde{f}-\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)\right\|_p\\ &\le\left\|\widetilde{f}-\widetilde{f_0}\right\|_p+\left\|\widetilde{f_0}-\frac{1}{n}\sum_{k=0}^{n-1}f_0(T^kx)\right\|_p+\left\|\frac{1}{n}\sum_{k=0}^{n-1}(f_0-f)(T^kx)\right\|_p.\end{aligned}$$

Now, $\widetilde{f}-\widetilde{f_0}=\widetilde{f-f_0}$ and hence,

$$\left\|\widetilde{f}-\widetilde{f_0}\right\|_p=\left\|\widetilde{f-f_0}\right\|_p\le\|f-f_0\|_p\le\frac{\varepsilon}{3},$$

and

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1}(f_0-f)(T^kx)\right\|_p\le\frac{1}{n}\sum_{k=0}^{n-1}\|f_0-f\|_p=\|f_0-f\|_p\le\frac{\varepsilon}{3}.$$

Therefore, for $n\ge N$,

$$\left\|\widetilde{f}-\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)\right\|_p<\varepsilon,$$

which implies that

$$\lim_{n\to\infty}\left\|\widetilde{f}-\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)\right\|_p=0.$$

We now prove the remainder of the statements in Corollary 3.9.

(ii)

$$\begin{aligned}\widetilde{f}(Tx) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(T^k(Tx)) \\ &= \lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=0}^{n} f(T^k x) - \frac{1}{n}f(x)\right) \\ &= \lim_{n\to\infty}\frac{n+1}{n}\frac{1}{n+1}\sum_{k=0}^{n} f(T^k x) - \lim_{n\to\infty}\frac{1}{n}f(x) \\ &= \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n} f(T^k x) \\ &= \widetilde{f}(x).\end{aligned}$$

(iii) If $f \in L^p(\mu)$, note that by *(ii)*, the sequence $\dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1} f(T^k x)$ converges to $\widetilde{f}$ in $L^1(\mu)$. Hence,

$$\int_X \widetilde{f}\,d\mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\int_X f(T^k x)\,d\mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\int_X f\,d\mu = \int_X f\,d\mu.$$

□

In Birkhoff Ergodic Theorem, suppose the limit $\widetilde{f}(x) = c$, where c is a constant. Then,

$$\int_X f\,d\mu = \int_X \widetilde{f}\,d\mu = c\mu(X).$$

That is,

$$c = \widetilde{f}(x) = \frac{1}{\mu(X)}\int_X f\,d\mu.$$

In other words, we see that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) = \frac{1}{\mu(X)}\int_X f\,d\mu.$$

The left hand side is the *time average* of f and the right hand side is the *space average* of f. This is what the physicists call the *ergodic hypothesis*, (the equality of the time and space averages of f).

Proposition 3.11 *Let T be an invertible measure preserving transformation of X, $f \in L^1(\mu)$ and let*

$$f_n^+(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) \quad f_n^-(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(T^{-k}x).$$

Then, $\widetilde{f}^+ = \lim_{n\to\infty} f_n^+$ and $\widetilde{f}^- = \lim_{n\to\infty} f_n^-$ exist and are equal almost everywhere, i.e., $\widetilde{f}^+ = \widetilde{f}^-$ -a.e.

Proof We first observe that

$$f_N^+ \circ T^{-(N-1)}(x) = \frac{1}{N}\sum_{k=0}^{N-1} f(T^k(T^{-(N-1)}x)) = \frac{1}{N}\sum_{k=0}^{N-1} f(T^{-k}x) = f_N^-(x).$$

Also, since $\widetilde{f}_N^+ \circ T = \widetilde{f}_N^+$ and $\widetilde{f}_N^- \circ T^{-1} = \widetilde{f}_N^-$, we get $\widetilde{f}_N^+ \circ T^{-1} = \widetilde{f}_N^+$ and hence, $\widetilde{f}_N^+ \circ T^{-k} = \widetilde{f}_N^+$ for all $k \in \mathbb{N}$. Therefore,

$$\begin{aligned}
\widetilde{f}_N^+(x) = \widetilde{f}_N^+ \circ T^{-(N-1)}(x) &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f_N^+(T^k(T^{-(N-1)}x)) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f_N^+ \circ T^{-(N-1)}(T^k x) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f_N^-(T^k x) \\
&= \widetilde{f}_N^-(x).
\end{aligned}$$

Hence $\widetilde{f}^+ = \lim_{n\to\infty} \widetilde{f}_n^+ = \lim_{n\to\infty} \widetilde{f}_n^- = \widetilde{f}^-$ (this holds because, $f_n \to f$ implies $\widetilde{f}_n \to \widetilde{f}$). □

Definition 3.12

1. A *measurable flow* in a measure space $(X, \mathfrak{M}, \mu)$ is a map $\tau : X \times \mathbb{R} \longrightarrow X$ that satisfies the following two conditions:
   (a) τ is measurable with respect to the product measure $\mu \times \lambda$ on $X \times \mathbb{R}$ and the measure μ on X. Here, λ is the Lebesgue measure on $\mathbb{R}$.
   (b) For $t \in \mathbb{R}$, the maps $\tau_t(x) := \tau(x, t)$ form a one-parameter group of transformations of X to itself with $\tau_0 =$ identity on X and $\tau_{t+s} = \tau_t \circ \tau_s$ for $t, s \in \mathbb{R}$.
2. A measurable flow τ_t is *measure preserving* or is μ-invariant if $\mu(\tau_t A) = \mu(A)$ for every $t \in \mathbb{R}$ and every $A \in \mathfrak{M}$.

Remark 3.13 If τ_t is a measure preserving flow on a finite measure space $(X, \mathfrak{M}, \mu)$ and if $f \in L^1(\mu)$, then the limits

$$f^+ = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(\tau_t x)\,dt \quad \text{and} \quad f^- = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(\tau_{-t} x)\,dt$$

exist and are equal for μ - a.e x.

Proof Let $F(x) = \int_0^1 f(\tau_t x)\,dt$. Since f and τ are measurable, $f \circ \tau(x,t) = f(\tau_t x)$ is measurable and by Fubini theorem $F(x) = \int_0^1 f(\tau_t x)\,dt$ is μ-measurable and is in $L^1(\mu)$ since $f \in L^1(\mu)$.

Now

$$\lim_{n\to\infty} \frac{1}{n}\int_0^n f(\tau_t x)\,dt = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} F(\tau_1^k(x))$$

(where $\tau_1(x) = \tau(x,1) : X \times \mathbb{R} \to X$) exists for μ a.e. x by Birkhoff ergodic theorem.

Let

$$\widetilde{f}(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} F(\tau_1^k(x)) = \lim_{n\to\infty} \frac{1}{n}\int_0^n f(\tau_t x)\,dt.$$

If $t \in \mathbb{R},\ t > 0$ is such that $n < t < n+1$ for $n \in \mathbb{N} \cup \{0\}$, then

$$\begin{aligned}
\left|\int_0^t f(\tau_t x)\,dt - \int_0^n f(\tau_t x)\,dt\right| &= \left|\int_n^t f(\tau_t x)\,dt\right| \\
&\le \left|\int_n^{n+1} f(\tau_t x)\,dt\right| \\
&\le \int_n^{n+1} |f(\tau_t x)|\,dt \\
&= \int_0^1 \left|f(\tau_1^n \circ \tau_t(x))\right|\,dt \\
&= \int_0^1 |f(\tau_t x)|\,dt,
\end{aligned}$$

where the last equality follows from Theorem 3.4.

Since

$$\frac{1}{n}\int_0^1 |f(\tau_t x)|\,dt \to 0 \ \text{ as } \ n \to \infty,$$

we have

$$\frac{1}{t}\left|\int_n^t f(\tau_t x)\,dt\right| \ \le \ \frac{1}{n}\left|\int_n^t f(\tau_t x)\,dt\right| \to 0 \ \text{ as } \ n \to \infty.$$

Since $t \to \infty$ as $n \to \infty$, we have

$$\frac{1}{t}\int_n^t f(\tau_t x)\,dt \to 0 \ \text{ as } \ t \to \infty,$$

and hence

$$\frac{1}{t}\int_0^t f(\tau_t x)\,dt \to \widetilde{f}(x) \text{ as } t \to \infty.$$

Now the remark follows by virtue of the preceding Proposition 3.11. □

Definition 3.14

1. Let $(X, \mathfrak{M}, \mu)$ be a probability space. If $A \in \mathfrak{M}$ and T is a measure preserving transformation of X, then A is said to be T-*invariant* if $\mu(T^{-1}A \Delta A) = 0$. A is said to be strictly T-invariant if $T^{-1}A = A$.
2. A measurable function $f : X \longrightarrow \mathbb{R}$ is T-*invariant* if $\mu(\{x : f(Tx) \neq f(x)\}) = 0$. f is *strictly* T-*invariant* if $f(Tx) = f(x)$ for all x.

The next two observations seek to bridge the divide between T-invariant and strictly T-invariant sets (or functions).

Lemma 3.15

1. *If $A \in \mathfrak{M}$ is a T-invariant set, then there is a strictly T-invariant set A_∞ such that $\mu(A_\infty) = \mu(A)$.*
2. *If f is a T-invariant function, then there is a strictly T-invariant function $\bar{f}$ such that $\bar{f}(x) = f(x)$ -a.e.*

Proof

1. Let

$$A_\infty = \bigcap_{n=0}^{\infty}\bigcup_{i=n}^{\infty} T^{-i}A.$$

   It is easy to check that $A_\infty \in \mathfrak{M}$, $T^{-1}A_\infty = A_\infty$ and $\mu(A_\infty) = \mu(A)$.
2. Let

$$A_f = \left\{x : f(T^k x) = f(x) \text{ for some } k \in \mathbb{N}\right\}.$$

   Clearly, A_f has measure 1, since the set $\{x : f(Tx) = f(x)\}$ is contained in A_f. Let

$$\bar{f}(x) = \begin{cases} f(y) & \text{if } y = T^k(x) \in A_f \text{ for some } k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

   It is easy to see that $\bar{f}$ is well-defined, strictly T-invariant and $\bar{f} = f$-a.e. □

Let us find out the conditions under which the limit $\widetilde{f}(x)$ in the ergodic theorem is constant a.e. for every $f \in L^1(\mu)$.

Suppose $\widetilde{f}(x) =$constant -a.e. for every $f \in L^1(\mu)$. Let $A \in \mathfrak{M}$ be a strictly T-invariant set and let χ_A be the characteristic function of A.

The ergodic theorem for χ_A implies $\int_X \widetilde{\chi}_A\,d\mu = \int_X \chi_A\,d\mu = \mu(A)$. Now

$$\widetilde{\chi}_A(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k x).$$

Since $A = T^{-1}A$, $Tx \in A$ if and only if $x \in T^{-1}A = A$ or $T^k x \in A$ if and only if $x \in T^{-k}A = A$ for $k \in \mathbb{N}$. Therefore,

$$\widetilde{\chi}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

By assumption, $\widetilde{\chi}_A(x)$ =constant -a.e. Therefore, $\widetilde{\chi}_A = 0$ or 1 -a.e. This implies $\mu(A) = 0$ or 1. That is, every T-invariant set has measure either 0 or 1.

Now, suppose on the contrary that if $A \in \mathfrak{M}$ is T-invariant then $\mu(A) = 0$ or 1. Let $f \in L^1(\mu)$ and let $\widetilde{f}(x)$ be the limit as in the ergodic theorem. By ergodic theorem, $\widetilde{f} \circ T = \widetilde{f}$ -a.e. on X.

Let

$$A(k,n) \;=\; \left\{ x \;:\; \frac{k}{2^n} \le \widetilde{f}(x) < \frac{k+1}{2^n} \right\} \;\text{ for }\; k \in \mathbb{Z},\; n \in \mathbb{N}.$$

Now

$$T^{-1}(A(k,n)) \Delta A(k,n) \subset \left\{ x \;:\; \widetilde{f} \circ T(x) \neq \widetilde{f}(x) \right\}.$$

Therefore,

$$\mu(T^{-1}(A(k,n)) \Delta A(k,n)) = 0$$

and hence, $A(k,n)$ is a T-invariant set and therefore $\mu(A(k,n)) = 0$ or 1.

Now, for a fixed $n \in \mathbb{N}$, $\bigcup_{k\in\mathbb{Z}} A(k,n) = X$ is a disjoint union. Therefore, for each $n \in \mathbb{N}$, there exists a unique $k_n \in \mathbb{Z}$ such that $\mu(A(k_n,n)) = 1$.

Let $Y = \bigcap_{n=1}^{\infty} A(k_n,n)$. Then $\mu(Y) = 1$ (because $\mu(Y^c) = 0$). Since $\widetilde{f}$ is constant on Y and $\mu(Y) = 1$, $\widetilde{f}$ is constant a.e on X.

Definition 3.16 A measure preserving transformation $T : X \longrightarrow X$, where $(X, \mathfrak{M}, \mu)$ is a probability measure space, is said to be *ergodic* if for every set $A \in \mathfrak{M}$ which is T-invariant, one has $\mu(A) = 0$ or 1.

Indeed we have shown that a measure preserving transformation T is ergodic if and only if every T-invariant function f is constant a.e. on X.

Proposition 3.17 *Let $(X, \mathfrak{M}, \mu)$ be a second countable probability measure space such that every non-empty open subset of X has positive measure. If $T : X \longrightarrow X$ is an ergodic transformation then*

$$\mu\left(\left\{x \;:\; \left\{T^n x \;:\; n \ge 0 \text{ is dense in } X\right\}\right\}\right) = 1.$$

That is, almost all points in X have dense orbits.

Proof Let $\{U_n\}_{n=1}^{\infty}$ be a basis for X. Let

$$Y \;=\; \left\{x \;:\; \left\{T^n x \;:\; n \geq 0 \text{ is dense in } X\right\}\right\}.$$

Clearly $x \notin Y$ if and only if there is a basic open set U_k such that $x \in \bigcap_{n=0}^{\infty}(X \setminus T^{-n}(U_k)) = P$, say. It is easy to see that $P \subset T^{-1}(P)$. Since T is measure preserving and $P \in \mathfrak{M}$, $\mu(T^{-1}P) = \mu(P)$. Therefore, $T^{-1}P \equiv P \pmod 0$ and hence, P is T-invariant. Also $U_n \cap P = \varnothing$ and since $\mu(U_k) > 0$, we must have $\mu(P) = 0$, which implies $\mu(P^c) = 1$. Ergo, P^c consists of points x whose T-orbits are dense in X. □

Example 3.18 Let $X = [0, 1)$ be equipped with the Lebesgue measure. If $c \in \mathbb{R}$, the map $T_c : X \longrightarrow X$ defined by

$$T_c(x) \;=\; x + c \pmod 1 \;=\; \{x + c\} \;\text{ i.e., fractional part of } x + c.$$

It is clear that T_c preserves the Lebesgue measure, and it is easy to see that if $c \in \mathbb{Q}$, then T_c is periodic and all orbits are finite having same cardinality. Therefore, T_c is not ergodic when c is rational.

Example 3.19 If X is the circle $\mathbb{S} = \{z \in \mathbb{C} \;:\; |z| = 1\}$ with the normalised Lebesgue measure, then $T : \mathbb{S} \longrightarrow \mathbb{S}$ defined as $T(z) = az$ is measure preserving, as can be easily verified. Then T is ergodic iff a is not a root of unity. For, suppose a is a root of unity, i.e., $a^p = 1$ for some $p \neq 0$. Then $f(z) = z^p$. Clearly $f \circ T = f$, but f is not constant a.e. Therefore, T is not ergodic.

Conversely, suppose a is not a root of unity and let $f(z) = \sum_{n=-\infty}^{\infty} b_n z^n$ be its Fourier expansion. Now, $f \circ T = f$ implies $\sum_{n=-\infty}^{\infty} b_n a^n z^n = \sum_{n=-\infty}^{\infty} b_n z^n$. Hence, $b_n(a^n - 1) = 0$. As $a^n \neq 1$, for any $n \neq 0$, we must have $b_n = 0$ for all $n \neq 0$. Consequently, it follows that f is constant a.e. and that T is ergodic. Alternatively, if $a = e^{2\pi i c n}$, then T is ergodic whenever c is irrational.

4 Geodesic Flows on Closed Surfaces

Let M be a compact or, more generally, a complete, smooth manifold endowed with a Riemannian metric g, and let SM denote the associated unit tangent bundle. That is,

$$SM \;=\; \{(x, v) \;:\; x \in M,\; v \text{ is a unit tangent vector to } M \text{ at } x\}.$$

For each $t \in \mathbb{R}$, consider the transformation $\phi^t : SM \longrightarrow SM$ defined as follows: Given $(x, v) \in SM$, let γ_v be the unique geodesic in M passing through the point $x \in M$ and with v as its tangent vector at x. Since M is a manifold which is complete,

γ_v is defined on all of $\mathbb{R}$. Moreover, given any two points $p, q \in M$ there exists a geodesic joining p and q that realises the distance between them. Now set

$$\phi^t(x, v) \;=\; (\gamma_v(t), \gamma_v'(t)). \tag{2}$$

It is easy to verify that ϕ^t as defined above for all $t \in \mathbb{R}$ constitutes a 1-parameter group of transformations, called the geodesic flow, and satisfies the following properties:

1. $\phi^t \circ \phi^s = \phi^{t+s} = \phi^{s+t} = \phi^s \circ \phi^t$ and $\phi^0 = \mathrm{Id}|_{SM}$.
2. ϕ^t is measure preserving where the measure under consideration is the Liouville measure given locally by the product of the Riemannian volume [form] on M, (i.e., $\sqrt{\det(g_{ij})}\, dx_1 \wedge \cdots \wedge dx_n$) - also called the Riemannian measure and the usual Lebesgue measure on the unit sphere.

It would be illuminating to look at a simple example of the geodesic flow.

Example 4.1 Suppose $M = \mathbb{S}^2$, the unit 2-sphere, then M admits a metric of constant positive curvature. Since all of its geodesics are great circles, it means that every orbit of the geodesic flow is periodic, and is therefore not ergodic.

Following up on the previous example, the question of ergodicity of the geodesic flow on closed surfaces of constant negative curvature is treated in the sequel.

The Gauss-Bonnet theorem suggests that a compact Riemann surface with genus ≥ 2 admits a Riemannian metric of constant negative curvature.

We shall initially see how to define such a metric on these surfaces. The universal cover of the surface is, in fact, the upper half plane $\mathbb{H}^2$, where $\mathbb{H}^2 = \{z \in \mathbb{C} \,:\, \mathrm{Im}(z) > 0\}$, equipped with the metric $ds = \dfrac{\sqrt{dx^2 + dy^2}}{y}$, which is a metric of constant negative curvature, called the *hyperbolic metric*. Therefore, we first discuss the geometry of the upper half plane.

4.1 Isometries and Geodesics of $\mathbb{H}^2$

Let $\gamma : I \longrightarrow \mathbb{H}^2$ be a piecewise differentiable path parametrised as

$$\gamma(t) \;=\; \left\{z(t) \;=\; x(t) + iy(t) \in \mathbb{H}^2 \;:\; t \in I\right\}, \quad \text{where } \; I = [0, 1].$$

Then, the hyperbolic length $l(\gamma)$ of the path is given by

$$l(\gamma) \;=\; \int_0^1 \frac{\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}}{y(t)}\, dt \;=\; \int_0^1 \frac{\left|\frac{dz}{dt}\right|}{y(t)}\, dt. \tag{3}$$

The hyperbolic distance $\rho_h(z, w)$ between any two points $z, w \in \mathbb{H}^2$ is given as $\rho_h(z, w) = \inf l(\gamma)$, where the infimum is taken over all piecewise differentiable paths γ joining z and w in $\mathbb{H}^2$.

A natural question is to look at the isometries of $\mathbb{H}^2$; i.e., transformations on $\mathbb{H}^2$ preserving the hyperbolic distance ρ_h defined above. This leads us to a particular group of matrices denoted as $\mathrm{PSL}(2, \mathbb{R})$.

In order to place the elements in $\mathrm{PSL}(2, \mathbb{R})$, we first look at the group of matrices $\mathrm{SL}(2, \mathbb{R})$ consisting of all 2×2 real matrices of the form

$$g \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\text{ where }\; \det(g) = 1. \tag{4}$$

Quite clearly, the above group of matrices assumes a correspondence with the group of all *fractional linear transformations* of $\mathbb{C}$ onto itself of the form

$$\left\{ z \longmapsto \frac{az+b}{cz+d} \;:\; ad - bc = 1;\; a, b, c, d \in \mathbb{R} \right\}$$

with the product of two such transformations being equivalent to the product of two corresponding matrices in $\mathrm{SL}(2, \mathbb{R})$ and the inverse of a given transformation corresponding to the inverse matrix.

However the correspondence is not 1-1, rather any such fractional linear transformation is represented by a pair of matrices $\pm g$. Ergo, the group of all fractional linear transformations, henceforth identified with $\mathrm{PSL}(2, \mathbb{R})$, is isomorphic to $\mathrm{SL}(2, \mathbb{R})/\pm I$, where I is the 2×2 identity matrix. The corresponding identity transformation in $\mathrm{PSL}(2, \mathbb{R})$ will be denoted by Id.

Remark 4.2 Note that $\mathrm{PSL}(2, \mathbb{R})$ contains all fractional linear transformations of the form $z \longmapsto \frac{az+b}{cz+d}$, where $ad - bc = \Delta > 0$, as dividing the numerator and denominator by $\sqrt{\Delta}$ gives a new matrix of determinant 1, but resulting in the same transformation on $\mathbb{H}^2$. In particular, $\mathrm{PSL}(2, \mathbb{R})$ contains transformations of the form $z \longmapsto az + b,\; a, b \in \mathbb{R},\; a > 0$ and those of the form $z \longmapsto \dfrac{-1}{z}$.

Remark 4.3 $\mathrm{PSL}(2, \mathbb{R})$ acts on $\mathbb{H}^2$ by homeomorphisms. In fact, $\mathrm{PSL}(2, \mathbb{R}) \subset \mathrm{Isom}(\mathbb{H}^2)$, the group of all isometries of $\mathbb{H}^2$ (i.e., transformations of $\mathbb{H}^2$ onto itself preserving the hyperbolic distance on $\mathbb{H}^2$).

Proof Firstly, any transformation of the form $z \longmapsto \dfrac{az+b}{cz+d}$ on $\mathbb{C}$ maps $\mathbb{H}^2$ onto itself. Given any $T \in \mathrm{PSL}(2, \mathbb{R})$, let $w = T(z) = \dfrac{az+b}{cz+d}$. Then,

$$w \;=\; \frac{(az+b)(cz+d)}{|cz+d|^2} \;=\; \frac{ac\,|z|^2 + adz + bc\bar{z} + bd}{|cz+d|^2}.$$

Hence, the imaginary part $\mathrm{Im}(w)$ of w is,

$$\mathrm{Im}(w) \;=\; \frac{w-\overline{w}}{2i} \;=\; \frac{z-\overline{z}}{2i\,|cz+d|^2} \;=\; \frac{\mathrm{Im}(z)}{|cz+d|^2}.$$

Therefore, $\mathrm{Im}(z) > 0 \iff \mathrm{Im}(w) > 0$. As T is continuous and its inverse exists, we conclude that T is a homeomorphism of $\mathbb{H}^2$ onto itself.

To show that $T \in \mathrm{PSL}(2,\mathbb{R})$ is an isometry of $\mathbb{H}^2$ onto itself, we show that if $\gamma : I \longrightarrow \mathbb{H}^2$ is a piecewise differentiable path in $\mathbb{H}^2$, then $l\,(T(\gamma)) = l(\gamma)$. Therefore, suppose $\gamma := z(t) = x(t) + iy(t)$, and $T(\gamma)$ is given by $w(t) = T(z(t)) = u(t) + iv(t)$. Now

$$\frac{dw}{dz} \;=\; \frac{a(cz+d)-c(az+b)}{(cz+d)^2} \;=\; \frac{1}{(cz+d)^2}.$$

Since $v = \dfrac{y}{|cz+d|^2}$, we have $\left|\dfrac{dw}{dz}\right| = \dfrac{v}{y}$. Therefore,

$$\begin{aligned} l(T(\gamma)) = \int_0^1 \frac{\left|\frac{dw}{dt}\right|}{v(t)}\,dt \;&=\; \int_0^1 \frac{\left|\frac{dw}{dz}\frac{dz}{dt}\right|}{v(t)}\,dt \\ = \int_0^1 \frac{\left|\frac{dw}{dz}\right|\left|\frac{dz}{dt}\right|}{v(t)}\,dt \;&=\; \int_0^1 \frac{\left|\frac{dz}{dt}\right|}{y(t)}\,dt \;=\; l(\gamma). \end{aligned}$$

□

It is a fact that isometries take geodesics to geodesics and hence any transformation in $\mathrm{PSL}(2,\mathbb{R})$ maps geodesics to geodesics. We now determine the geodesics on the hyperbolic plane.

Theorem 4.4 *The geodesics in $\mathbb{H}^2$ are semicircles and straight lines orthogonal to the real axis.*

Proof Let $z_1, z_2 \in \mathbb{H}^2$. First suppose $z_1 = ia$ and $z_2 = ib$ with $b > a$ which are two points on the imaginary axis. If $\gamma : [0,1] \longrightarrow \mathbb{H}^2$ is any path joining ia to ib, with $\gamma(t) = x(t) + iy(t)$, then

$$l(\gamma) = \int_0^1 \frac{\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}}{y(t)}\,dt \geq \int_0^1 \frac{\left|\frac{dy}{dt}\right|}{y(t)}\,dt \geq \int_a^b \frac{dy}{y} \geq \ln\frac{b}{a}.$$

It is easy to verify that the equality in the above expression is realised by the hyperbolic length of the segment of the y-axis joining ia to ib which is of length $\ln\dfrac{b}{a}$ and hence the geodesic joining the points ia and ib is the segment of the imaginary axis between them.

If $z_1, z_2 \in \mathbb{H}^2$ are arbitrary, let L be the unique Euclidean semi-circle or straight line orthogonal to the real axis passing through z_1 and z_2, then there exists

a transformation in $\mathrm{PSL}(2,\mathbb{R})$ which maps L into the imaginary axis. The transformation $T(z) = \dfrac{-1}{z-a}$ takes a to ∞ and b to $\dfrac{1}{b-a}$ (>0), and the transformation $S(z) = z - \dfrac{1}{b-a} = z - c$ takes ∞ to ∞ and c to 0. Thus,

$$S \circ T = \begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -a \end{pmatrix} = \begin{pmatrix} -c & -1-ac \\ 1 & -a \end{pmatrix}$$

is the transformation in $\mathrm{PSL}(2,\mathbb{R})$ that takes (a, b) to $(\infty, 0)$. Since each element of $\mathrm{PSL}(2,\mathbb{R})$ is an isometry of $\mathbb{H}^2$ and segments of the imaginary axis are geodesics, we conclude that the geodesic joining z_1 and z_2 is the segment of L joining them. □

Since $\mathrm{PSL}(2,\mathbb{R})$ acts by isometries on $\mathbb{H}^2$, it acts on the unit tangent bundle $S\mathbb{H}^2$ as

$$g(z,\zeta) = (g(z), D_z g(\zeta)) = \left(g(z), \frac{1}{(cz+d)^2}\right),$$

where $z \in \mathbb{H}^2$, $\zeta \in T_z\mathbb{H}^2$ such that $\|\zeta\| = 1$ and $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2,\mathbb{R})$.

Lemma 4.5 *The action of* $\mathrm{PSL}(2,\mathbb{R})$ *on* $S\mathbb{H}^2$ *is transitive and free, i.e., all isotropy groups are trivial.*

Proof Let $z_0 = i$ and ζ_0 be the unit tangent vector at z_0 pointing in the positive direction of the imaginary axis. Let $(z,\zeta) \in S\mathbb{H}^2$ and σ be the positive imaginary half axis starting from z_0. Let L be the unique geodesic determined by (z,ζ). Let $g \in \mathrm{PSL}(2,\mathbb{R})$ be the transformation taking σ to L, i.e., $g(\sigma) = L$, with $g(z_0) = z$. Since transformations of $\mathrm{PSL}(2,\mathbb{R})$ have positive determinant, they preserve orientation and hence the condition that $D_{z_0} g(\zeta_0) = \zeta$ forces g to be unique; we will, therefore, denote it by $g_{z\zeta}$. □

Remark 4.6 In the above lemma, taking $(z,\zeta) \in S\mathbb{H}^2$ to $g_{z\zeta} \in \mathrm{PSL}(2,\mathbb{R})$, sets up a bijection F between $S\mathbb{H}^2$ and $\mathrm{PSL}(2,\mathbb{R})$, and is easily seen to be a diffeomorphism.

Let $z_0 = i$ and ζ_0 be as in the proof of Lemma 4.5. Given an arbitrary $(z,\zeta) \in S\mathbb{H}^2$, let $g_{z\zeta}$ be the unique element of $\mathrm{PSL}(2,\mathbb{R})$ (which exists by virtue of the lemma) that takes (z_0,ζ_0) to (z,ζ) in $S\mathbb{H}^2$. The uniqueness of the element $g_{z\zeta}$ shows that the diffeomorphism F intertwines the action of $\mathrm{PSL}(2,\mathbb{R})$ on $S\mathbb{H}^2$ with the left multiplication in the group. That is,

$$g((z,\zeta)) = g \cdot g_{z\zeta} \quad \forall g \in \mathrm{PSL}(2,\mathbb{R}).$$

Proposition 4.7 *The geodesic flow on* $S\mathbb{H}^2$ *corresponds to the flow on the group* $\mathrm{PSL}(2,\mathbb{R})$ *given by the right translation*

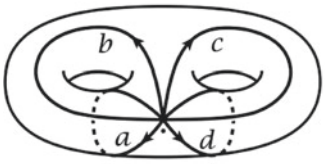
$$g \longmapsto g \cdot g_t, \quad \textit{where} \quad g_t = \begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{\frac{-t}{2}}, \end{pmatrix} \quad \forall t \in \mathbb{R}.$$

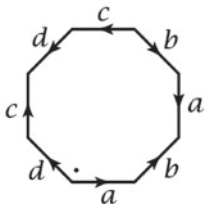
Proof It is clear that $\phi^t(z_0, \zeta_0) = g_t(z_0, \zeta_0)$, where ϕ^t is the geodesic flow. Therefore, for $(z, \zeta) \in S\mathbb{H}^2$,

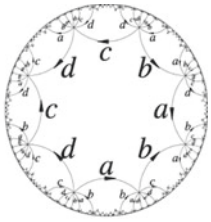
$$\phi^t(z, \zeta) = \phi^t\left(g_{z\zeta}(z_0, \zeta_0)\right) = g_{z\zeta}\left(\phi^t(z_0, \zeta_0)\right) = g_{z\zeta}\left(g_t(z_0, \zeta_0)\right) = g_{z\zeta} g_t.$$

The second equality is a result of the fact that the action of $\mathrm{PSL}(2, \mathbb{R})$ on $\mathbb{H}^2$ is by isometries, and hence takes geodesics to geodesics as described in the proof of Lemma 4.5. □

Let Σ be a compact Riemann surface of genus $g \geq 2$. Then Σ has $\mathbb{H}^2$ as its universal cover, i.e., if $\Gamma = \pi_1(\Sigma)$, the fundamental group of Σ, then Γ acts freely and discontinuously on $\mathbb{H}^2$ by deck transformations. Consequently, Γ can be identified with a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R})$ such that the quotient space $\Sigma = \mathbb{H}^2/\Gamma$ is compact. Further Σ is a Riemannian manifold with constant negative curvature -1 with respect to the metric induced from $\mathbb{H}^2$ via the quotient map. The pictures in this page roughly serve to illustrate this procedure.







Proposition 4.8 *The identification of* $S\mathbb{H}^2$ *with* $\mathrm{PSL}(2, \mathbb{R})$ *induces an identification* $S\left(\mathbb{H}^2/\Gamma\right) \cong \Gamma\backslash\mathrm{PSL}(2, \mathbb{R})$. *The geodesic flow on* $S\Sigma$ *corresponds to the flow*

$$\Gamma \setminus \mathrm{PSL}(2, \mathbb{R}) \longrightarrow \Gamma\backslash\mathrm{PSL}(2, \mathbb{R}), \quad \Gamma g \longmapsto \Gamma g g_t,$$

where $g_t = \begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{\frac{-t}{2}} \end{pmatrix}$.

Proof Since $(z, \zeta) \longmapsto g_{z\zeta}$ intertwines the action of $\mathrm{PSL}(2, \mathbb{R})$, the proof follows from the previous proposition and is left as an exercise to the reader. □

4.2 Hopf's Proof of Ergodicity

In this section, we sketch a proof of the ergodicity of the geodesic flow g_t on $\Gamma\backslash\mathrm{PSL}(2, \mathbb{R})$ that was originally presented by E. Hopf [9]. In this context, we introduce the notion of horocycles, some of whose illustrative examples are the lines parallel to the x-axis in $\mathbb{H}^2$. As we shall soon discover, horocycles have a very special role in the study of the dynamics of the geodesic flow.

Lines parallel to the x-axis can also be viewed as orbits of points in $\mathbb{H}^2$ under the action of the 1-parameter subgroup of $\mathrm{PSL}(2, \mathbb{R})$ consisting of matrices of the form

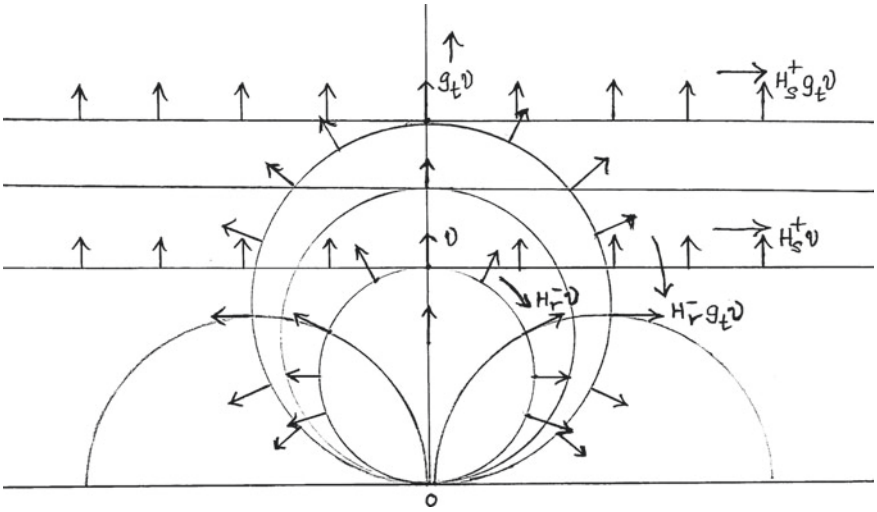


Fig. 1 Geodesic and horocycle flows

$H_s^+ = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$; that is, transformations of the form $z \longmapsto z + s$. Being orthogonal to the lines parallel to the y-axis in $\mathbb{H}^2$, it turns out that their images, under a typical element of PSL(2, $\mathbb{R}$) taking ∞ to a point x_0 on the x-axis, are the Euclidean circles in $\mathbb{H}^2$ tangent to the x-axis at the point x_0.

Moving a step further, and using the identification of PSL(2, $\mathbb{R}$) with $S\mathbb{H}^2$, we see that the 1-parameter subgroup H_s^+, of PSL(2, $\mathbb{R}$), defines a measure preserving flow on $S\mathbb{H}^2$. In a similar fashion, we observe that the 1-parameter subgroup $H_r^- = \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix}$ of PSL(2, $\mathbb{R}$) also defines a measure preserving flow on $S\mathbb{H}^2$. The flow H_s^+ is termed *the stable horocycle flow* while H_r^- is termed *the unstable horocycle flow*.

The next figure serves to illustrate the orbits of a vector $v \in S\mathbb{H}^2$ under the dynamics of the two horocycle flows, in relation to the geodesic flow.

The two horocycle flows determine vector fields on $S\mathbb{H}^2$ which are linearly independent, i.e., at any given point of $S\mathbb{H}^2$, the tangent vectors of the corresponding vector fields are linearly independent and hence, together with the tangent vector given by geodesic flow vector field, span the tangent space to $S\mathbb{H}^2$ at that point.

4.2.1 A Historical Interlude

Eberhard Hopf exploited the interrelation between the stable and unstable horocycle flows and the geodesic flow in his proof. Historically it was G.A. Hedlund [7] who, in 1934, first proved that the geodesic flow on closed surfaces of constant negative curvature is ergodic (which was called metric transitivity at that time). In 1936, E.

Hopf gave another proof of ergodicity in the case considered by Hedlund. Hedlund was also the first to recognize the importance of the close relationship between horocycle and geodesic flows. Later, in 1939, Hedlund proved [8] stronger properties (like mixing) for geodesic flow on surfaces of finite area and constant negative curvature. Erogdicity was extended to arbitrary dimensions for manifolds of constant negative curvature by Hopf in 1939. In the same paper [9], Hopf also proved that the geodesic flow is ergodic for a surface of finite area and of variable negative curvature under the restriction that the curvature and its first derivatives are bounded in absolute value (Fig. 1).

Gelfand and Fomin, in 1952 [5], provided the next impetus by proving the stronger property of mixing for the case of manifolds of higher dimension and constant negative curvature. Their approach and method was generalised by Mautner in 1957 [11] to prove ergodicity of the geodesic flow on locally symmetric spaces of negative curvature and arbitrary dimensions.

However the question remained open in the case of variable curvature in arbitrary dimension until 1960s when the work of Anosov and Sinai [2] led Anosov to prove ergodicity for closed manifolds of negative curvature and arbitrary dimension [1]. The approach adopted in the work of Anosov and Sinai enabled Anosov to overcome the difficulty faced by Hopf, and Anosov proved ergodicity for manifolds of finite volume and variable negative curvature under exactly the same hypothesis considered by Hopf in 1939 [9], namely when the covariant derivative of the curvature tensor is bounded in absolute value.

Remark 4.9 For manifolds of finite volume and variable negative curvature without the boundedness assumption on the first derivatives of curvature, to the best of our knowledge, the question of ergodicity is still an outstanding open problem (even for surfaces!).

Resuming the sketch of Hopf's proof, let $f : S\Sigma \longrightarrow \mathbb{R}$ be a continuous function with compact support where Σ is a surface of genus $g \geq 2$ with the hyperbolic metric. Note that as a consequence of Theorem 2.39, it suffices to consider continuous functions with compact support. We will show that f is constant a.e. when f is g_t-invariant.

For the three smooth flows g_t, H_s^+ and H_r^- on $\mathrm{PSL}(2, \mathbb{R})$, a routine computation shows that

$$H_s^+ g_t \;=\; g_t H^+_{e^{-t}s} \;\text{ and }\; H_r^- g_t \;=\; g_t H^-_{e^{-t}r}.$$

From this, it follows that

$$f(x H_s^+ g_t) \;=\; f(x g_t H^+_{e^{-t}s}) \;\text{ and }\; f(x H_r^- g_t) \;=\; f(x g_t H^-_{e^{-t}r}).$$

Uniform continuity of f then implies that

$$\lim_{t\to\infty} \left(f(x H_s^+ g_t) - f(x g_t)\right) \;=\; \lim_{t\to\infty} \left(f(x g_t H^+_{e^{-t}s}) - f(x g_t)\right) \;=\; 0$$

and

$$\lim_{t\to\infty}\left(f(xH_r^-g_t)-f(xg_t)\right) \;=\; \lim_{t\to\infty}\left(f(xg_tH_{e^{-t}r}^-)-f(xg_t)\right) \;=\; 0.$$

Therefore,

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau \left(f(xg_t)-f(xH_s^+g_t)\right)\,dt \;=\; 0.$$

Similarly,

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau \left(f(xg_{-t})-f(xH_r^-g_{-t})\right)\,dt \;=\; 0.$$

With the notation introduced in an earlier remark in this chapter, we note that $\widetilde{f}^+(xH_s^+)$ and $\widetilde{f}^-(xH_r^-)$ exist whenever $\widetilde{f}^+(x)$ and $\widetilde{f}^-(x)$ exist. Further, we conclude from the above that $\widetilde{f}^+(x)=\widetilde{f}^+(xH_s^+)$ and $\widetilde{f}^-(x)=\widetilde{f}^-(xH_r^-)$, and are equal a.e.

Let $x_0\in S\Sigma$. We will construct an open neighbourhood of x_0 as follows. Let $\delta_1,\delta_2,\delta_3>0$ be sufficiently small. Construct a smooth curve $\gamma_{\delta_1}(x_0)$ through x_0 by defining

$$\gamma_{\delta_1}(x_0) \;=\; \left\{x_0H_r^- \;:\; |r|<\delta_1\right\}$$

and then construct an open smooth surface $\sigma_{\delta_1,\delta_2}(x_0)$ by defining

$$\begin{aligned}\sigma_{\delta_1,\delta_2}(x_0) &= \left\{x_0H_r^-g_t \;:\; |r|<\delta_1,\,|t|<\delta_2\right\}\\ &= \bigcup_{|t|<\delta_2}\left(\gamma_{\delta_1}(x_0)\right)g_t.\end{aligned}$$

Finally, construct an open neighbourhood $U_{\delta_1,\delta_2,\delta_3}(x_0)$ by

$$U_{\delta_1,\delta_2,\delta_3}(x_0) \;=\; \bigcup_{|s|<\delta_3}\left(\sigma_{\delta_1,\delta_2}(x_0)\right)H_s^+.$$

It follows from the smoothness of the corresponding vector fields that for sufficiently small $\delta_1,\delta_2,\delta_3$, the surfaces $\left(\sigma_{\delta_1,\delta_2}(x_0)\right)H_s^+$ are disjoint for distinct s with $|s|<\delta_3$ and for the point

$$x \;=\; x_0H_r^-g_tH_s^+ \;\in\; U_{\delta_1,\delta_2,\delta_3}(x_0),$$

the numbers r,t,s are smooth coordinates in $U_{\delta_1,\delta_2,\delta_3}(x_0)$. In fact, as x_0 varies over a compact set on $S\Sigma$, all of $\delta_1,\delta_2,\delta_3$ can be chosen to be independent of x_0. Now, the Liouville measure on $S\Sigma$ induces conditional measures on each of the surfaces $\left(\sigma_{\delta_1,\delta_2}(x_0)\right)H_s^+$, for all s and invoking Fubini's theorem shows that for a.e. $y\in\sigma_{\delta_1,\delta_2}(x_0)$ (with respect to the induced conditional measure), one has $\widetilde{f}^+(y)=\widetilde{f}^-(y)$; and this holds for x_0 a.e. in $S\Sigma$(with respect to μ).

We will now show that $\widetilde{f}(x)$ is constant for $x(= x_0 H_r^- g_t H_s^+)$ a.e. in $U_{\delta_1,\delta_2,\delta_3}(x_0)$. To this end, let

$$\widetilde{U} = \Big\{ x \in U_{\delta_1,\delta_2,\delta_3}(x_0) \ : \ \widetilde{f}^+(x) \text{ exists and } \\ \text{for } y = x_0 H_r^- g_t \in \sigma_{\delta_1,\delta_2}(x_0),\ \widetilde{f}^+(y) = \widetilde{f}^-(y) \Big\}.$$

Since the vector fields are smooth, it follows from Fubini's theorem that $\widetilde{U}$ has full measure in $U_{\delta_1,\delta_2,\delta_3}(x_0)$. Further, if $x_1, x_2 \in \widetilde{U}$, with $x_1 = x_0 N_{r_1}^- g_{t_1} N_{s_1}^+$ and $x_2 = x_0 N_{r_2}^- g_{t_2} N_{s_2}^+$, and if y_1, y_2, z_1, z_2 denote $x_0 N_{r_1}^- g_{t_1}$, $x_0 N_{r_2}^- g_{t_2}$, $x_0 N_{r_1}^-$ and $x_0 N_{r_2}^-$ respectively, then we have,

$$\begin{aligned} \widetilde{f}^+(x_1) = \widetilde{f}^+(y_1) &= \widetilde{f}^-(y_1) = \widetilde{f}^-(z_1) \\ = \widetilde{f}^-(z_2) &= \widetilde{f}^-(y_2) = \widetilde{f}^+(y_2) = \widetilde{f}^+(x_2). \end{aligned}$$

Thus $\widetilde{f}^+$ is constant in $\widetilde{U}$, i.e., $\widetilde{f}^+$ is constant a.e. in $U_{\delta_1,\delta_2,\delta_3}(x_0)$, which proves the ergodicity of g_t.

References

1. Anosov, D. V. (1967). Geodesic flows on closed Riemannian manifolds of negative curvature. *Proceedings of the Steklov Institute of Mathematics, 90*, 235.
2. Anosov, D. V., & Sinai, Y. G. (1967). Some smooth ergodic systems. *Russian Mathematical Surveys, 22*(5), 107–172.
3. Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America, 17*, 656–660.
4. Caratheodory, C. (1914). Über das lineare Mass von Punktmengen -eine Verallgemeinerung des Lïgenbegriffs, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (pp. 404–426).
5. Gel'fand, I. M., & Fomin, S. V. E. (1952). Geodesic flows on manifolds of constant negative curvature. *Uspekhi Matematicheskikh Nauk, 7*, 118–137.
6. Hausdorff, F. (1918). Dimension und äusseres mass. *Mathematische Annalen, 79*, 157–179.
7. Hedlund, G. A. (1934). On the metrical transitivity of the geodesics on closed surfaces of constant negative curvature. *Annals of Mathematics, 35*, 787–808.
8. Hedlund, G. A. (1939). Fuchsian groups and mixtures. *Annals of Mathematics, 40*, 370–383.
9. Hopf, E. (1939). Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. *Ber. Verh. Sachs. Akad. Wiss. Leipzig, 91*, 261–304.
10. Lebesgue, H. (1904). *Leçons sur l'intégration et la recherche des fonctions primitives*. Paris: Gauthier-Villars.
11. Mautner, F. I. (1957). Geodesic flows on symmetric Riemann spaces. *Annals of Mathematics, 65*(2), 416–431.
12. Vitali, G. (1905). Sul problema della misura dei gruppi di punti di una retta, Bologna, Tip. Gamberini e Parmeggiani.
13. von Neumann, J. (1932). Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 18*, 70–82.

Symbolic Dynamics

Siddhartha Bhattacharya

1 Introduction

In this chapter, we will study a class of topological dynamical systems known as symbolic dynamical systems. These systems play an important role in coding theory, combinatorial dynamics and theory of cellular automata. In Sect. 2, we introduce the basic concepts associated with such systems. In Sect. 3, we introduce the notion of entropy. In Sect. 4, we compute the measure theoretic entropy of Bernoulli shifts. In Sect. 5, we consider a class of symbolic dynamical systems related to tiling spaces, and prove a result due to M. Szegedy that asserts that any translational tiling of $\mathbb{Z}^d$ by a finite set F is periodic when $|F|$ is prime. The last section is devoted to an algebraic dynamical system known as 3-dot system. Using the concept of directional homoclinic groups we show that $\mathbb{Z}^2$-actions on symbolic spaces can exhibit strong rigidity property.

2 Basic Concepts

In this section, we review some basic concepts of symbolic dynamics (see [5] for a comprehensive introduction).

Definition 2.1 Let G be a discrete group.

1. A *topological G-space* is a compact topological space X together with a continuous action σ of G on X. In other words, σ is a continuous map from $G \times X$ to X that satisfies the properties of a group action.

S. Bhattacharya (✉)
School of Mathematics, Tata Institute for Fundamental Research, Mumbai, India
e-mail: siddhart@math.tifr.res.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_4

Notation 2.2 For any $g \in G$, the map $x \longmapsto g \cdot x = \sigma(g, x)$ will be denoted by $\sigma(g)$.

2. If (X, σ) and (Y, ρ) are topological G-spaces, a map $f : X \longrightarrow Y$ is said to be *G-equivariant* if $f \circ \sigma(g) = \rho(g) \circ f$ for all $g \in G$.
3. A topological G-space (Y, ρ) is said to be a *factor* of a topological G-space (X, σ) if there exists a surjective G-equivariant map from X to Y.
4. Two topological G-spaces (X, σ) and (Y, ρ) are *topologically conjugate* if there exists a G-equivariant homeomorphism from X to Y.

Let $A = \{1, \ldots, k\}$ be a finite set and let $A^{\mathbb{Z}}$ be the set of all functions from $\mathbb{Z}$ to A. The set $Y = A^{\mathbb{Z}}$ can also be viewed as the collection of all bi-infinite sequences taking values in A. For any $a \in A^{\mathbb{Z}}$, $\{a_i\}_{i \in \mathbb{Z}}$ will denote the corresponding bi-infinite sequence. Let d denote the discrete metric on A, i.e., $d(x, y) = 1$ if $x \neq y$ and $d(x, y) = 0$ if $x = y$. We define a metric d_Y on Y by

$$d_Y(a, b) \;=\; \sum_{i=-\infty}^{\infty} \frac{d(a_i, b_i)}{2^{|i|+1}}.$$

We note that $d_Y(a, b)$ is small if and only if there exists a large $N > 0$ such that $a_i = b_i$ for all $i \in [-N, N]$. Hence, d_Y induces the product topology on $Y = A^{\mathbb{Z}}$ with cylinder sets as basic open subsets (the choice of the metric is not relevant here as long as it induces the product topology).
Let $T : Y \longrightarrow Y$ be the shift map defined by $T(a)_i = a_{i+1}$. It is easy to see that (Y, d) is a compact metric space and $T : Y \longrightarrow Y$ is a self homeomorphism.

Definition 2.3 If $X \subset A^{\mathbb{Z}}$ is a closed shift invariant subset and T is the restriction of the shift map to X then (X, T) is called a *symbolic dynamical system*.

Example 2.4 $X = A^{\mathbb{Z}}$ and T is the shift map.

Example 2.5 Suppose we only have two symbols, i.e., $A = \{0, 1\}$. Let $X = \{a \in A^{\mathbb{Z}} \;:\; \text{there are no two consecutive 0's}\}$.

Example 2.6 We fix finite sets A and $E \subset A \times A$. Let G denote the directed graph with A as the set of vertices and E as the set of edges. We define $X_G = \{a \in A^{\mathbb{Z}} \;:\; (a_i, a_{i+1}) \in E \;\forall i\}$. The dynamical system $(X_G, T|_{X_G})$ is called the *topological Markov chain* corresponding to G. Note that Example 2.5 can be seen as a special case where $A = \{0, 1\}$ and $E = \{(1, 0), (1, 1), (0, 1)\}$.

Example 2.7 Suppose $A = \{0, 1\}$ and X is the set of all bi-infinite sequences in $\{0, 1\}$ such that between any two consecutive 1's there are even number of 0's. Then it is easy to verify that X is closed and shift-invariant.

Example 2.8 For any finite set A we define $L(A) = \bigcup_{n=1}^{\infty} A^n$. The set $L(A)$ can be viewed as the collection of all finite words with A as the alphabet set. For any $S \subset L(A)$, we define

$$X_S \;=\; \left\{a \in A^{\mathbb{Z}} \;:\; s \text{ does not occur in } A \;\forall s \in S\right\}.$$

Clearly, X_S is a closed shift-invariant subset of $A^{\mathbb{Z}}$.

Definition 2.9 Suppose $X \subset A^{\mathbb{Z}}$ is a closed shift-invariant subset and σ is the shift action of $\mathbb{Z}$ on X. Then (X, σ) is called a *subshift of finite type* if $X = X_S$ for some finite set $S \subset L(A)$.

Definition 2.10 Suppose $Y \subset A^{\mathbb{Z}}$ is a closed shift-invariant subset such that the shift action of $\mathbb{Z}$ on Y is a factor of a subshift of finite type $X \subset A^{\mathbb{Z}}$. Then the shift action on Y is called a *sofic shift*.

Example 2.11 With three symbols, suppose $A = \{0, 1, 2\}$ and $E = \{(1, 1), (1, 0), (2, 1), (0, 2), (2, 0)\}$. Let $X \subset A^{\mathbb{Z}}$ be the topological Markov chain associated with (A, E). Let $\phi : A \longrightarrow \{0, 1\}$ denote the map defined by $\phi(1) = 1$ and $\phi(0) = \phi(2) = 0$. Then, ϕ induces a continuous shift equivariant map from X to $A^{\mathbb{Z}}$. It is easy to see that the image of ϕ is the system described in Example 2.7. Hence the system described in Example 2.7 is a sofic shift.

Let A be a finite set. Fix $k \geq 1$, and choose a map $\theta : A^{2k+1} \longrightarrow A$. Such maps are called *block codes*. Any block code θ induces a map $\overline{\theta} : A^{\mathbb{Z}} \longrightarrow A^{\mathbb{Z}}$ defined by $\overline{\theta}(x)_i = \theta(x_{i-k}, \ldots, x_{i+k})$. The map $\overline{\theta}$ is called the *sliding block code* corresponding to θ.

Example 2.12 Let $A = \{0, 1\}$ and q be the continuous shift-equivariant map from $A^{\mathbb{Z}}$ to $A^{\mathbb{Z}}$ defined by $q(x)_i = x_{i-1} + x_i + x_{i+1} \pmod 2$. Then $q = \overline{\theta}$, where $\theta : A^3 \longrightarrow A$ is the block code defined by $\theta(a, b, c)_i = a + b + c \pmod 2$.

It is easy to see that for any block code θ, $\overline{\theta}$ is a continuous shift-equivariant map from $A^{\mathbb{Z}}$ to $A^{\mathbb{Z}}$. The following result, known as the Curtis-Hedlund theorem, shows that the converse is also true.

Theorem 2.13 *Suppose A is a finite set and $f : A^{\mathbb{Z}} \longrightarrow A^{\mathbb{Z}}$ is a continuous shift-equivariant map. Then there exists $k \geq 1$ and a block code $\theta : A^{2k+1} \longrightarrow A$ such that $f = \overline{\theta}$.*

Proof Since $A^{\mathbb{Z}}$ is compact, f is uniformly continuous, we choose a positive δ such that $d(f(x), f(y)) < \dfrac{1}{2}$ whenever $d(x, y) < \delta$. Since

$$d(f(x), f(y)) \;=\; \sum \frac{d(f(x)_i, f(y)_i)}{2^i},$$

it follows that $f(x)_0 = f(y)_0$ whenever $d(f(x), f(y)) < \dfrac{1}{2}$. We choose k such that $\displaystyle\sum_{|i|>k} \frac{1}{2^k} < \delta$. Then. $f(x)_0 = f(y)_0$ whenever $x_i = y_i$ for all i with $|i| \leq k$. This shows that there is a block code $\theta : A^{2k+1} \longrightarrow A$ such that $f(x)_0 = \theta(x_{-k}, \ldots, x_k)$. Since f is also shift-equivariant, we deduce that $f = \overline{\theta}$. □

3 Entropy

We will now introduce a dynamical invariant called topological entropy for symbolic dynamical systems. We will need the following elementary result about sequences of real numbers.

Proposition 3.1 *Let $\{a_i\}$ be a sequence of non-negative real numbers such that $a_{m+n} \leq a_m + a_n$ for all m and n. Then $\lim_{n\to\infty} \frac{a_n}{n}$ exists.*

Proof Set $c = \inf_n \frac{a_n}{n}$. For any $\epsilon > 0$, we choose n such that

$$\left|\frac{a_n}{n} - c\right| \;<\; \epsilon.$$

Let $D = \max\{a_1, \ldots, a_n\}$. Let $m \geq n$ be any positive integer. We write $m = kn + j$, where $0 \leq j \leq n-1$. Now,

$$\frac{a_m}{m} \;\leq\; \frac{ka_n + a_j}{kn + j} \;\leq\; c + \epsilon + \frac{D}{m}.$$

This shows that $\frac{a_m}{m} \leq c + 2\epsilon$ as $m \to \infty$. Since ϵ is arbitrary, we conclude that $\frac{a_m}{m} \to c$ as $m \to \infty$. □

For $m \leq n$, let $[m, n]$ denote the set $\{m, \ldots, n\}$. For any closed shift invariant subset $X \subset A^{\mathbb{Z}}$ and a finite set $S \subset \mathbb{Z}$, let π_S denote the projection map from $A^{\mathbb{Z}}$ to A^S. For $k \geq 1$, let B_k denote the set $\pi_{[0,k-1]}(X)$. The set B_k can also be described as the set of all blocks of length k that occurs in elements of X. Since X is shift invariant it follows that $\pi_{[0,\,n-1]}(X) = \pi_{[m,\,m+n-1]}(X)$ for all m and n. Since there is a natural injective map from $\pi_{[0,\,m+n-1]}(X)$ to $\pi_{[0,\,m-1]}(X) \times \pi_{[m,\,m+n-1]}(X)$, we deduce that $|B_{m+n}| \leq |B_m| \times |B_n|$. We define

$$h(X) \;=\; \lim_{k\to\infty} \frac{\log(|B_k|)}{k}.$$

The number $h(X)$ is called the *entropy* of the shift action of $\mathbb{Z}$ on X. By the previous proposition it is well defined.

Example 3.2 Suppose $X = A^{\mathbb{Z}}$. In this case $B_n = A^n$ and $|B_n| = |A|^n$. Hence, the entropy of the corresponding shift action is $\log|A|$.

Example 3.3 Suppose $X = \{a \in \{0,1\}^{\mathbb{Z}} \;:\; \text{there are no two consecutive 1's}\}$. Let T denote the 2×2 adjacency matrix of the associated graph. Then, $T_{11} = T_{12} = T_{21} = 1$ and $T_{22} = 0$. Hence T has two distinct eigenvalues $\frac{\sqrt{5} \pm 1}{2}$. It is easy to see that $|B_n|$ is the sum of entries of T^{n-1}. This implies that

$$h(X) \;=\; \lim_{n\to\infty} \frac{\log|B_n|}{n} \;=\; \log\left(\frac{\sqrt{5}+1}{2}\right).$$

We now show that topological entropy is invariant under topological conjugacy.

Theorem 3.4 *Let A be a finite set and for $i = 1, 2$, let X_i be a closed shift invariant subset of $A^{\mathbb{Z}}$ such the corresponding shift actions of $\mathbb{Z}$ are topologically conjugate. Then $h(X_1) = h(X_2)$.*

Proof Let f be a topological conjugacy between these two shift actions. From Curtis-Hedlund theorem, it follows that there exists $k \geq 1$, and a map $\theta : A^{2k+1} \longrightarrow A$ such that f is the sliding block code corresponding to θ. Hence for any $i \leq j$, the elements $f(x)_i, \ldots, f(x)_j$ are determined by the elements $x_{i-k}, \ldots, x_{j+k}$. In particular, $|B_n(X_2)| \leq |B_{n+2k}(X_1)|$. Taking logarithms, dividing by n, and letting $n \to \infty$, we see that $h(X_1) \geq h(X_2)$. Similarly, we can show that $h(X_2) \geq h(X_1)$. □

Our next task is to define the notion of entropy for a more general class of dynamical systems.

Definition 3.5 Let L be an abelian semigroup with the property that $x + x = x$ for all $x \in L$. A *norm* on L is a map $\|\cdot\|$ from L to $\mathbb{R}^+$ satisfying

$$\|x\| \;\leq\; \|x+y\| \;\leq\; \|x\| + \|y\| \quad \forall x, y \in L.$$

A *normed lattice* is an abelian semigroup L together with a norm map $\|\cdot\| : L \longrightarrow \mathbb{R}^+$.

Example 3.6 Let S be a set and let L be the collection of all finite subsets of S. For $A, B \in L$ set $A + B = A \cup B$, and $\|A\| = |A|$, the cardinality of A.

Example 3.7 Let V be a vector space and let L be the collection of all finite dimensional subspaces of V. For $X, Y \in L$, define $X + Y$ to be the smallest subspace containing X and Y, and set $\|X\| = \dim(X)$.

Example 3.8 Let X be a compact topological space. An open cover C of X is called *saturated* if for any two open subsets U and V of X with $U \in C$ and $V \subset U$, we have $V \in C$. Let L be the collection of all saturated open covers of X. For $C,\ C^{'} \in L$, we define $C + C^{'}$ to be the collection of all open subsets that belong to both C and $C^{'}$. It is easy to see that $C + C^{'}$ is an element of L. For any $C \in L$, we define $\|C\| = \log(n_C)$, where n_C is the smallest cardinality of a subcover of C.

Notation 3.9 For $x, y \in L$, we say $x \leq y$ if $x + y = y$.

It is easy to see that the above notation defines a partial order on L.

Definition 3.10 If $T : L \longrightarrow L^{'}$ is a map between normed lattices then T is called an *isometry* if $T(x + y) = T(x) + T(y)$ and $\|T(x)\| = \|x\|$ for all x, y. Clearly, the collection of all normed lattices form a category with isometries as *morphisms*.

If $T : L \longrightarrow L$ is an isometry, then we define

$$\|\cdot\|_T : L \longrightarrow \mathbb{R}^+ \text{ by } \|x\|_T = \lim_{n\to\infty} \frac{1}{n}(x + Tx + \cdots + T^{n-1}x).$$

Proposition 3.11 *The map $\|\cdot\|_T$ is well defined and it is a norm on L. Furthermore, it satisfies the following two properties:*

1. *$\|x\|_T \leq \|x\|$ for all x in L;*
2. *Both T and $I + T$ are isometries with respect to $\|\cdot\|_T$.*

Proof Fix any $x \in L$ and define a sequence $\{a_n\}$ by

$$a_n = \left\|x + Tx + \cdots + T^{n-1}x\right\|.$$

Since T is an isometry, applying the sub-additivity of the norm, we see that $a_{m+n} \leq a_m + a_n$ for all $m, n \geq 1$. From Proposition 3.1, we deduce that $\|\cdot\|_T$ is well defined. It is easy to see that $\|\cdot\|_T$ is a norm and satisfies property *1*. Since $x + x = x$ in L, we obtain

$$\sum_{i=0}^{n-1} T^i(x + Tx) = \sum_{i=0}^{n} T^i(x),$$

which proves the second property. □

Definition 3.12 For any isometry $T : L \longrightarrow L$, we define the *entropy* of T by

$$h(T) = \sup\{\|x\|_T : x \in L\}.$$

Definition 3.13 Let (X, μ) be a measure space with $\mu(X) = 1$. A *partition* $P = \{P_1, \ldots, P_m\}$ of X is a finite collection of pairwise disjoint, non-empty, measurable subsets of X such that $\bigcup P_i = X$.

Notation 3.14 Let L_X be the set of all partitions of X. For $P, Q \in L_X$, we define

$$P + Q = \left\{P_i \cap Q_j : P_i \in P,\ Q_j \in Q \text{ and } P_i \cap Q_j \neq \varnothing\right\}.$$

It is easy to see that L_X becomes an abelian semigroup and $P + P = P$ for all P. For $P = \{P_1, \ldots, P_m\} \in L_X$, we set

$$\|P\| = -\sum_{i=1}^{m} \mu(P_i)\ \log_2(\mu(P_i)).$$

Proposition 3.15 *L_X is a normed lattice with respect to the above norm.*

Proof Choose $P = \{P_1, \ldots, P_m\}$ and $Q = \{Q_1, \ldots, Q_n\}$ in L. Set $p_i = \mu(P_i)$, $q_j = \mu(Q_j)$ and $r_{ij} = \mu(P_i \cap Q_j)$. Now,

$$\|P+Q\| - \|P\| \;=\; \sum p_i \log p_i - \sum r_{ij} \log r_{ij} \;=\; -\sum r_{ij}\left(\log r_{ij} - \log p_i\right).$$

Since log is an increasing function, this shows that $\|P+Q\| \geq \|P\|$.
Define $\phi : [0,1] \longrightarrow \mathbb{R}$ by $\phi(0) = 0$ and $\phi(x) = -x \log x$ if $x > 0$. Since $\phi''(x) = -\dfrac{1}{x} < 0$ in $(0,1)$, it follows that ϕ is a concave function. Put $c_{ij} = \dfrac{r_{ij}}{p_i}$ if $p_i > 0$ and 0 otherwise. Observe that $\|P+Q\| - \|P\| = \sum p_i \phi(c_{ij})$. Since ϕ is concave, we deduce that

$$\|P+Q\| - \|P\| \;\leq\; \sum_j \phi\left(\sum_i p_i c_{ij}\right) \;=\; \sum_j \phi(q_j) \;=\; \|Q\|.$$

□

If $T : (X,\mu) \longrightarrow (Y,\nu)$ is a measure preserving map then, we define a map $T^* : L_Y \longrightarrow L_X$ by

$$T^*(P) \;=\; \left\{T^{-1}(P_1), \ldots, T^{-1}(P_n)\right\}.$$

It is easy to see that T^* is an isometry. Moreover, the correspondence $X \longmapsto L_X$ and $T \longmapsto T^*$ gives us a contravariant functor from the category of probability spaces to the category of normed lattices. If T is a measure preserving map from (X,μ) to itself then we define $h(T) = h(T^*)$, where T^* is the isometry of L_X induced by T. The number $h(T)$ is called the *entropy* of T. Clearly, entropy is a measurable conjugacy invariant.
Suppose X is a compact topological space and T is a homeomorphism of X. As in the Example 3.8, let L denote the collection of all saturated open covers of X. For any $C \in L$, we define $T^*(C) = \left\{T^{-1}(U) \;:\; U \in C\right\}$. It is easy to see that $T^*(C) \in L$ for all $C \in L$ and T^* is an isometry of L. The number $h(T^*)$ is called the *topological entropy* of T. It is a topological conjugacy invariant. In the special case when (X,T) is a one-dimensional shift, this coincides with the more explicit definition presented earlier.

4 Computations of Entropy

In this section, we compute the entropy of Bernoulli shifts and translations on tori. If X is a set and $\mathcal{A}$ is a collection of subsets of X then by $\sigma(\mathcal{A})$ we denote the smallest σ-algebra on X that contains $\mathcal{A}$. We begin with the following approximation lemma.

Lemma 4.1 *Suppose $(X,\mathcal{B},\mu)$ is a probability space and suppose $\mathcal{A} \subset \mathcal{B}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{B}$. Then for any $P \in L_X$ and $\epsilon > 0$, there exist a partition $P_1 \subset \mathcal{A}$ and $Q \in L_X$ with $\|Q\| < \epsilon$ such that $P \leq P_1 + Q$.*

Proof We first consider the case when P has only two elements, i.e., $P = \{B, B^c\}$ for some measurable set B. Note that $x \log x \to 0$ as $x \to 0$ or $x \to 1$. Hence, we can

find $\delta > 0$ such that $\mu(E) < \delta$ implies $\|\{E, E^c\}\| < \epsilon$. As $\sigma(\mathcal{A}) = \mathcal{B}$, we can find $A \in \mathcal{A}$ such that $\mu(F) < \delta$, where $F = (B \setminus A) \cup (A \setminus B)$. Define $P_1 = \{A, A^c\}$ and $Q = \{F, F^c\}$. It is easy to see that P_1 and Q have the required properties.

Now suppose $P = \{B_1, \ldots, B_n\}$. For $1 \leq i \leq n$, define $P^i = \left\{B_i, B_i^c\right\}$. Find $P_1^i,\ Q^i$ as above with $\left\|Q^i\right\| < \dfrac{\epsilon}{n}$ and put $P_1 = \sum P^i$ and $Q = \sum Q^i$. □

We note the following consequence of the previous lemma.

Proposition 4.2 *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \longrightarrow X$ be a measure preserving map. Suppose $\mathcal{A}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{B}$. Then $h(T) = \sup\{\|P\|_T \ :\ P \subset \mathcal{A}\}$.*

Proof Fix $\epsilon > 0$ and choose P' such that $h(T) \leq \left\|P'\right\|_T + \epsilon$. Applying the previous lemma, find P_1 and Q such that $P' \leq P_1 + Q,\ P_1 \subset \mathcal{A}$ and $\|Q\| < \epsilon$. Since $\|Q\|_T \leq \|Q\|$, it follows that

$$h(T) \ \leq\ \|P_1\|_T + \|Q\|_T + \epsilon \ =\ \|P_1\|_T + 2\epsilon.$$

As ϵ is arbitrary, this proves the proposition. □

Definition 4.3 Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \longrightarrow X$ be an invertible measure preserving map. A partition P is said to be a *generator* if $\mathcal{B}$ is the smallest σ-algebra that is invariant under the $\mathbb{Z}$-action generated by T and contains $\{P_1, \ldots, P_n\}$.

Theorem 4.4 *If P is a generator, then $h(T) = \|P\|_T$.*

Proof For any partition P, let $A(P)$ denote the collection of all subsets which can be expressed as unions of elements of P. It is easy to verify that $A(P)$ is a finite algebra and $Q \leq P$ if and only if $Q \subset A(P)$. We define an algebra A_∞ by

$$A_n \ =\ A\left(\sum_{-n}^{n} T^{*i}\right), \quad A_\infty \ =\ \bigcup_{n=1}^{\infty} A_n.$$

Note that A_∞ is the smallest T-invariant algebra containing P. Hence. $\sigma\,(A_\infty) = \mathcal{B}$. If a partition Q is contained in A_∞ then $Q \subset A_n$ for some n. Hence,

$$\|Q\|_T \ \leq\ \left\|\sum_{i=-n}^{n} T^{*i} P\right\|_T \ =\ \left\|\left(I + T^*\right)^{2n+1}(P)\right\|_T \ =\ \|P\|_T\,.$$

From the previous lemma it then follows that $h(T) = \|P\|_T$. □

Definition 4.5 Let (X, μ) be a probability space and let $P, Q \in L_X$. Then P and Q are said to be *independent* if $\mu(P_i \cap Q_j) = \mu(P_i)\mu(Q_j)$ for all i and j.

It is easy to see that if P and Q are independent then $\|P + Q\| = \|P\| + \|Q\|$.

4.1 Entropy of Shifts

Let $Y = \{y_1, \ldots, y_n\}$ be a finite set and let ν be a probability measure on Y. Let $(X, \mathcal{B}, \mu) = (Y, \nu)^{\mathbb{Z}}$ and let $T : X \longrightarrow X$ be the shift map. We define a partition $P = \{P_1, \ldots, P_n\}$ of X by

$$P_i \;=\; \{x \in X \;:\; x(0) = y_i\}\,.$$

Let $\mathcal{A}$ be the smallest T-invariant σ-algebra containing P. Since $P \subset \mathcal{A}$, the co-ordinate projection corresponding to 0th co-ordinate is a $\mathcal{A}$-measurable map. Since $\mathcal{A}$ is T-invariant, all co-ordinate projections are measurable. Hence $\mathcal{A} = \mathcal{B}$, i.e., P is a generator. We observe that for any k, the partitions $P + \cdots + T^{*k-1}P$ and $T^{*k}P$ are independent. Applying induction on k, we see that $\left\| \sum_{i=0}^{k-1} T^{*i} P \right\| = k \, \| P \|$. Hence, $h(T) = \| P \|_T = \| P \|$. In the special case, when ν is the uniform measure on Y, $h(T) = \log n$.

Proposition 4.6 *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \longrightarrow X$ be a measure preserving map.*

1. *$h(T^n) = nh(T)$ for all $n \geq 1$.*
2. *If T is invertible then $h(T^{-1}) = h(T)$.*

Proof We will prove the statements for any lattice isometry $T : L \longrightarrow L$.

1. Fix $x \in L$ and put $y = x + Tx + \cdots + T^{n-1}x$. Note that

$$\sum_{i=0}^{k-1} T^{in} x \;\leq\; \sum_{i=0}^{nk-1} T^{i} x \;=\; \sum_{i=0}^{k-1} T^{ni} y.$$

   This shows that $\| x \|_{T^n} \leq n \, \| x \|_T = \| y \|_{T^n}$. Since x is arbitrary, we conclude that $h(T^n) = nh(T)$.
2. If T is invertible then for any $x \in L$,

$$\left\| \sum_{i=0}^{k-1} T^{-i} x \right\| \;=\; \left\| T^{1-k} \left(\sum_{i=0}^{k-1} T^{i} x \right) \right\| \;=\; \left\| \sum_{i=0}^{k-1} T^{i} x \right\|.$$

   Hence $\| x \|_T = \| x \|_{T^{-1}}$ for all x and $h(T) = h(T^{-1})$.

□

For $i = 1, 2$, let $(X_i, \mathcal{B}_i, \mu_i)$ be a probability space and let $T_i : X_i \longrightarrow X_i$ be a measure preserving map. We define $T_1 \times T_2 : X_1 \times X_2 \longrightarrow X_1 \times X_2$ by $(T_1 \times T_2)\,(x, y) = (T_1 x, T_2 y)$. It is easy to see that $T_1 \times T_2$ preserves the measure $\mu_1 \times \mu_2$.

Proposition 4.7 $h\,(T_1 \times T_2) \;=\; h\,(T_1) + h\,(T_2)$.

Proof For $i = 1, 2$, let π^i denote the projection map from $X_1 \times X_2$ to X_i. Since π^i is measure-preserving, π^i_* is an isometry from L_{X_i} to $L_{X_1 \times X_2}$. It is easy to see that $(T_1 \times T_2)^k_* \pi^i_*(P) = \pi^i_*(T^k_{i*} P)$ for any P in L_{X_i}. We note that for any $P \in L_{X_1}$ and $Q \in L_{X_2}$, the partitions $\pi^1_*(P)$ and $\pi^2_*(Q)$ are independent. Hence, for arbitrary P and Q,

$$\left\| \pi^1_*(P) + \pi^2_*(Q) \right\|_{T_1 \times T_2} \;=\; \|P\|_{T_1} + \|Q\|_{T_2}\,.$$

This implies that $h\,(T_1 \times T_2) \geq h\,(T_1) + h\,(T_2)$.
Let $\mathcal{A}$ denote the algebra of all subsets of $X_1 \times X_2$ that can be expressed as a finite union of measurable rectangles. If R is a partition of $X_1 \times X_2$ such that $R \subset \mathcal{A}$, then we can find $P \in L_{X_1}$ and $Q \in L_{X_2}$ such that $R \leq \pi^1_*(P) + \pi^2_*(Q)$. Since $\sigma(\mathcal{A})$ is the product σ-algebra on $X_1 \times X_2$, applying Proposition 4.2 and the above equality, we see that $h\,(T_1 \times T_2) \leq h\,(T_1) + h\,(T_2)$. □

Lemma 4.8 *Let $P = \{P_1, \ldots, P_n\}$ be a partition of a probability space (X, μ). Then $\|P\| \leq \log n$.*

Proof Put $p_i = \mu(P_i)$. Then $\|P\| = \sum p_i \log\left(\dfrac{1}{p_i}\right)$. As $x \longmapsto \log x$ is a concave function, we see that $\|P\| \leq \log\left(\sum p_i \cdot \dfrac{1}{p_i}\right) = \log n$. □

4.2 Entropy of Translations

Let $n \geq 1$ and let $\theta = (\theta_1, \ldots, \theta_n)$ be an element of the n-torus $\mathbb{T}^n$. Let $T : \mathbb{T}^n \longrightarrow \mathbb{T}^n$ denote the map $x \longmapsto \theta \cdot x$. We claim that $h(T) = 0$. Note that $T = T_1 \times \cdots \times T_n$, where $T_i : \mathbb{T} \longrightarrow \mathbb{T}$ is the translation by θ_i. By Proposition 4.7, $h(T) = \sum h(T_i)$. Hence, without loss of generality, we may assume that $n = 1$.

Case 1. $\theta^k = 1$ for some k. Since $P + P = P$ for all P, it follows that $\|P\|_{\mathrm{Id}} = 0$ for all P, i.e., $h(\mathrm{Id}) = 0$. Since $T^k = \mathrm{Id}$, applying Proposition 4.6, we see that $h(T) = 0$.

Case 2. θ is not a root of unity. We consider the partition $P = \{P_1, P_2\}$ where

$$P_1 \;=\; \{z \,:\, 0 \leq z < \pi\}\,, \quad P_2 \;=\; \{z \,:\, \pi \leq z < 2\pi\}\,.$$

Since $\{\theta^n \;:\; n \in \mathbb{Z}\}$ is dense in $\mathbb{T}$, it follows that P is a generator for T. Hence, $h(T) = \|P\|_T$. Note that for any $k \geq 1$, the partition $P + \cdots + T^{k-1}_* P$ has $2k$ sets. By the previous lemma, $\|P\|_T \leq \lim\limits_{k \to \infty} \dfrac{\log 2k}{k} = 0$, which proves the claim.

5 Tilings

For any finite set A and $d \geq 1$, the compact space $A^{\mathbb{Z}^d}$ admits a shift action of $\mathbb{Z}^d$. If $d > 1$, and X is a closed shift invariant subset of $A^{\mathbb{Z}^d}$ then the restriction of the shift action to X is called a *higher-dimensional shift*. In this section, we consider a class of such systems that arises from tilings of $\mathbb{Z}^d$.

Notation 5.1 For $d \geq 1$, let A, B and C be subsets of $\mathbb{Z}^d$. We will write $A \oplus B = C$ if every element of C can be uniquely expressed as $a + b$, with $a \in A$ and $b \in B$.

Definition 5.2 If $F \subset \mathbb{Z}^d$ is a finite set, then a *tiling* of $\mathbb{Z}^d$ by F is a subset C of $\mathbb{Z}^d$ satisfying $F \oplus C = \mathbb{Z}^d$.

It is easy to see that F tiles $\mathbb{Z}^d$ if and only if $\mathbb{Z}^d$ can be written as a disjoint union of translates of F.

Definition 5.3 A set $E \subset \mathbb{Z}^d$ is said to be *periodic* if there exists a finite index subgroup $\Lambda \subset \mathbb{Z}^d$ such that $E + \Lambda = E$.

Let $F = \{g_1, \ldots, g_n\}$ be a finite subset of $\mathbb{Z}^d$. We equip $\{0, 1\}^{\mathbb{Z}^d}$ with the product topology and define $X(F) \subset \{0, 1\}^{\mathbb{Z}^d}$ by

$$X(F) \;=\; \left\{\mathbf{1}_C \;:\; F \oplus C = \mathbb{Z}^d\right\}.$$

It is easy to see that $x \in X(F)$ if and only if for each $g \in \mathbb{Z}^d$ there exists exactly one $g' \in g - F$ such that $x(g') = 1$. This shows that $X(F)$ is a closed subset of the compact space $\{0, 1\}^{\mathbb{Z}^d}$. Moreover, $X(F)$ is invariant under the shift action of $\mathbb{Z}^d$. The space $X(F)$ can be viewed as the space of all tilings of F. It is non-empty if and only if $\mathbb{Z}^d$ can be tiled by F.

Example 5.4 Suppose $d = 2$, and $F = \{(0, 0), (1, 0), (-1, 0), (0, -1)\}$. If an element $(m, n) \in \mathbb{Z}^2$ corresponds to the square $(m,\ m + 1] \times (n,\ n + 1] \in \mathbb{R}^2$, then the set F corresponds to a T-shaped set in $\mathbb{R}^2$. It is easy to verify that there is a unique $C \in \mathbb{Z}^2$ such that $(0, 0) \in C$ and $F \oplus C = \mathbb{Z}^2$. This implies that any tiling of F is a translate of C by an element of $-F$. In particular, F admits exactly 4 tilings, and all tilings of F are periodic.

Example 5.5 Suppose $d = 2$, and $F = \{(0, 0), (1, 0)\}$. Then the tilings of $\mathbb{Z}^2$ by F are in bijective correspondence with the tilings of the plane by 2×1 rectangle. We fix an element $\mathbf{1}_C$ of $X(F)$ and define a map $h_C : \mathbb{Z} \longrightarrow \{0, 1\}$ by $h_C(i) = \mathbf{1}_C((0,\ i))$. It is easy to see that the $C \longmapsto h_C$ is a bijective correspondence between $X(F)$ and the set of all maps from $\mathbb{Z}$ to $\{0, 1\}$. Hence $X(F)$ can be identified with the compact space $\{0, 1\}^{\mathbb{Z}}$. The shift action of $\mathbb{Z}^2$ on $X(F) = \{0, 1\}^{\mathbb{Z}}$ can be explicitly described. The element $(0,\ 1)$ acts by the shift map on $\{0, 1\}^{\mathbb{Z}}$, and the element $(1, 0)$ acts by flipping the symbols.

We note that in the previous example the space $X(F)$ is infinite but every element of $X(F)$ is periodic in the direction of $(1, 0)$. The following example shows that this need not be true in general.

Example 5.6 Suppose $d = 2$, and $F = \{(0, 0), (2, 0), (0, 2), (2, 2)\}$. We define $E_1 = \{(m, n) : m \text{ is even}\}$ and $E_2 = \{(m, n) : m \text{ is odd}\}$. We note that the tilings of E_1 by F are in bijection with the tilings of $\mathbb{Z}^2$ by $F' = \{(0, 0), (1, 0), (0, 2), (1, 2)\}$. Hence as in the previous example, we can find $C_1 \subset \mathbb{Z}^2$ such that $C_1 \oplus F = E_1$ and C_1 is periodic in the direction of $(1, 0)$ but not in the direction of $(0, 1)$. Similarly we can find C_2 such that $C_2 \oplus F = E_2$ and C_2 is periodic in the direction of $(0, 1)$ but not in the direction of $(1, 0)$. If we define C to be the disjoint union of C_1 and C_2 then $C \in X(F)$ and it is not periodic in any direction.

The following conjecture is due to Lagarias and Wang [6]:

Conjecture 5.7 (Periodic tiling conjecture) *Suppose $d \geq 1$ and $F \subset \mathbb{Z}^d$ is a finite set such that $F \oplus C = \mathbb{Z}^d$ for some $C \in \mathbb{Z}^d$. Then there exists a periodic set $E \subset \mathbb{Z}^d$ such that $F \oplus E = \mathbb{Z}^d$.*

The following proposition shows that a stronger version is true in the 1-dimensional case.

Proposition 5.8 *Let F and C be subsets of $\mathbb{Z}$ such that F is finite and $F \oplus C = \mathbb{Z}$. Then C is periodic.*

Proof Without loss of generality we may assume that $0 \in F$. Let k denote the diameter of F. From the condition $F \oplus C = \mathbb{Z}$, we deduce that for any $i \in \mathbb{Z}$, $\sum_{j \in F} \mathbf{1}_C(i + j) = 1$. Let B denote the block $(0, \ldots, k - 1)$. Suppose C and C' are two tilings of $\mathbb{Z}$ by F such that the restrictions of $\mathbf{1}_C$ and $\mathbf{1}_{C'}$ to B are equal. Then the above condition implies that $\mathbf{1}_C(k) = \mathbf{1}_{C'}(k)$. By taking $i = 1, 2, \ldots$ and applying this argument repeatedly we see that $\mathbf{1}_C(j) = \mathbf{1}_{C'}(j)$ for all $j \geq 0$. A similar argument shows that $\mathbf{1}_C(j) = \mathbf{1}_{C'}(j)$ for all $j \leq 0$. Combining these two observations, we deduce that $C = C'$. Since B is a block of length k, this implies that there are only finitely many $C \subset \mathbb{Z}$ such that $F \oplus C = \mathbb{Z}$. As any translate of a tiling is again a tiling, we conclude that every tiling of $\mathbb{Z}$ by F is periodic. □

Definition 5.9 A subset $F \subset \mathbb{Z}^d$ is sdid to be *non-degenerate* if $0 \in F$ and the elements of F generate a finite index subgroup of $\mathbb{Z}^d$.

The following theorem due to M. Szegedy (see [8]) describes the tilings of a non-degenerate set F when the number of elements of F is prime.

Theorem 5.10 *Let F, C be subsets of $\mathbb{Z}^d$ such that F is finite and $F \oplus C = \mathbb{Z}^d$. If F is non-degenerate and $|F|$ is a prime number then, C is periodic.*

Proof Let M_d denote the set of all functions from $\mathbb{Z}^d$ to $\mathbb{R}$. There is a natural action θ of $\mathbb{Z}^d$ on M_d defined by

$$\theta(g)(f)(x) \;=\; f(x-g) \;\; \forall x, g \in \mathbb{Z}^d.$$

It is easy to see that $F \oplus C = \mathbb{Z}^d$ if and only if $\sum_{g\in F} \theta(g)(\mathbf{1}_C) = \mathbf{1}_{\mathbb{Z}^d}$. If $F = \{g_1, \ldots, g_p\}$, where p is a prime number, then this shows that

$$\left(\sum_{g\in F} \theta(g)\right)^p (\mathbf{1}_C) \;=\; \left(\sum_{g\in F} \theta(g)\right)^{p-1} (\mathbf{1}_{\mathbb{Z}^d}) \;=\; p^{p-1}\mathbf{1}_{\mathbb{Z}^d}.$$

On the other hand,

$$\begin{aligned}\left(\sum_{g\in F} \theta(g)\right)^p (\mathbf{1}_C) &= \left(\theta(g_1)^p + \cdots + \theta(g_p)^p\right)(\mathbf{1}_C) \\ &= \left(\theta(pg_1) + \cdots + \theta(pg_p)\right)(\mathbf{1}_C) \pmod p.\end{aligned}$$

Hence C satisfies the equation

$$\sum_{g\in F} \theta(pg)\,(\mathbf{1}_C) \;=\; 0 \pmod p.$$

Now let w be an arbitrary element of $\mathbb{Z}^d$. Then $\theta(pg)\,(\mathbf{1}_C)\,(w) \in \{0, 1\}$ for all $g \in F$. Since their sum is divisible by p, we conclude that either $\theta(pg)\,(\mathbf{1}_C)\,(w) = 1$ for all $g \in F$ or $\theta(pg)\,(\mathbf{1}_C)\,(w) = 0$ for all $g \in F$. In particular, $\theta(pg)\,(\mathbf{1}_C) = \mathbf{1}_C$ for all $g \in F$. Hence $\mathbf{1}_C$ is invariant under the translations by elements of the subgroup generated by $\left\{pg_i - pg_j \;:\; g_i, g_j \in F\right\}$. Since F is non-degenerate, it follows that this subgroup has finite index. This implies that C is periodic. □

Let F be a finite non-degenerate subset of $\mathbb{Z}^d$ such that $|F|$ is a prime number and let H denote the subgroup generated by F. We pick a finite set $E \subset \mathbb{Z}^d$ such that E contains exactly one element from each coset of H. It is easy to see that subsets of E are in bijective correspondence with the H-invariant subsets of $\mathbb{Z}^d$. The proof of the previous theorem shows that $X(F)$ is finite and has at most $2^{|\mathbb{Z}^d/H|}$ elements.

6 3-Dot Shifts

Let $\mathbb{Z}_2$ denote the group $\mathbb{Z}/2\mathbb{Z}$ and let Y denote the set $\mathbb{Z}_2^{\mathbb{Z}^2}$. It is easy to see that Y is a compact abelian group with respect to pointwise addition and the product topology. We define the shift action σ of $\mathbb{Z}^2$ on Y by $(\sigma(n)(x))\,(m) = x(m+n)$ for

all $m, n \in \mathbb{Z}^2$. It is easy to see that $\sigma(n)$ is an automorphism of Y for all $n \in \mathbb{Z}^2$. Let $R_d = \mathbb{Z}_2[\mathbb{Z}^d]$ denote the group-ring of $\mathbb{Z}^d$ with coefficients in $\mathbb{Z}_2$. Alternatively, one can identify R_d with $\mathbb{Z}_2[U_1^{\pm}, \ldots, U_d^{\pm}]$, the ring of Laurent polynomials in d commuting variables with coefficients in $\mathbb{Z}_2$. For any $f = \sum_{n \in \mathbb{Z}^d} c_n u^n$ and $y \in Y$, we define $f \cdot y \in Y$ by

$$f \cdot y \;=\; \sum_{n \in \mathbb{Z}^d} c_n \sigma(n)(x).$$

It is easy to see that Y becomes a module with respect to this operation. For any ideal $I \subset R_d$, let $Y(I) \subset Y$ denote the closed subgroup defined by $Y(I) = \{y \in Y \;:\; f \cdot y = 0 \; \forall f \in I\}$. It is easy to see that $Y(I)$ is a σ-invariant subgroup for any I. Using Pontryagin duality, one can show that this correspondence between closed shift invariant subgroups of Y and ideals of R_d is bijective.
In this section, we will look at a specific higher dimensional shift that arises this way. Let $d = 2$, $f = 1 + U_1 + U_2$ and $I \subset R_2$ be the principal ideal generated by f, i.e., $I = fR_2$. Then $X = Y(I)$ is called the 3-*dot system*. We note that if τ denotes the automorphism of Y defined by $\tau = \sigma(1, 0) + \sigma(0, 1) + I$, then $X = \{x \in Y \;:\; \tau(x) = 0\}$. This system was first introduced by F. Ledrappier in order to study mixing properties of algebraic dynamical systems (see [4, 7] for more details). Using Pontryagin duality theory, one can show that (X, σ) is irreducible in the sense that every proper closed shift invariant subgroup of X is finite.

Definition 6.1 Suppose G and H are abelian topological groups. A continuous map $\phi : G \longrightarrow H$ is called *affine* if there exists a continuous homomorphism $\theta : G \longrightarrow H$ and $b \in H$ such that $\phi(g) = \theta(g) + b$ for all $g \in G$.

For any $f : G \longrightarrow H$, we define $\widehat{f} : G \times G \longrightarrow H$ by $\widehat{f}(x, y) = f(x + y) - f(x) - f(y) + f(0)$.

Lemma 6.2 *A continuous map f is affine if and only if $\widehat{f} = 0$.*

Proof It is easy to see that if f is affine then $\widehat{f}$ vanishes. Conversely suppose $\widehat{f}$ is identically zero. Set $b = f(0)$ and define $\theta : G \longrightarrow H$ by $\theta(x) = f(x) - f(0)$. Clearly $f(x) = \theta(x) + b$ for all $x \in G$. Moreover, for any $x, y \in G$, $\theta(x + y) - \theta(x) - \theta(y) = \widehat{f}(x, y) = 0$. This proves the given assertion. □

Definition 6.3 Suppose $d \geq 1$ and σ is a continuous action of $\mathbb{Z}^d$ on a compact metric space X. For $x, y \in X$, the pair (x, y) is called *homoclinic* if $d(\sigma(m)(x), \sigma(m)(y)) \longrightarrow 0$ as $\|m\| \to \infty$.

Example 6.4 Suppose $d = 1$, $X = \mathbb{T}$ and σ is given by a rotation. Since every rotation is an isometry, (x, y) is a homoclinic pair if and only if $x = y$.

Example 6.5 Suppose $d = 1$ and σ is the shift action on $\{0, 1\}^{\mathbb{Z}}$. Then (x, y) is a homoclinic pair if and only if $x_i = y_i$ for all but finitely many i's.

If X is a compact abelian group then (x, y) is a homoclinic pair if and only if $(x - y, 0)$ is a homoclinic pair. If σ is a continuous action of $\mathbb{Z}^d$ on a compact abelian group X by automorphisms of X, then we define

$$\Delta_\sigma(X) \;=\; \{x \in X \,:\, \sigma(n)(x) \to 0 \text{ as } \|n\| \to \infty\}\,.$$

It is easy to see that $\Delta_\sigma(X)$ is a subgroup of X. It is called the *homoclinic group* of the action σ.

Lemma 6.6 *Let (X, σ) denote the 3-dot system. Then, $\Delta_\sigma(X) = \{0\}$.*

Proof As $[\sigma(1,0) + \sigma(0,1) + \sigma(0,0)](x) = 0$ for all $x \in X$ and every element of X has order 2, it follows that for all $k \geq 1$,

$$[\sigma(1,0) + \sigma(0,1) + \sigma(0,0)]^{2^k} \;=\; \sigma(2^k, 0) + \sigma(0, 2^k) + \sigma(0,0) \;=\; 0.$$

This implies that for any $x \in X$ and $(m, n) \in \mathbb{Z}^2$, $x(m + 2^k, n) + x(m, n + 2^k) + x(m, n) = 0$. If x is homoclinic to 0 then the first two terms vanish for large k, and hence $x = 0$. □

Definition 6.7 Let X be a compact abelian group and σ, an action of $\mathbb{Z}^d$ on X by continuous automorphisms. Suppose v is an element of the unit sphere $S^{d-1} \subset \mathbb{R}^d$. An element $x \in X$ is called *v-homoclinic* if $\sigma(g)(x) \to 0$ as $\langle v,\ g\rangle \to \infty$.

For any $v \in S^{d-1}$, the collection of all v-homoclinic points are denoted by $\Delta_v(\sigma)$. It is easy to see that $\Delta_v(\sigma)$ is a subgroup of X. As we will see shortly, these groups can be non-trivial, even when the homoclinic group of the action σ is trivial.
Suppose σ is the shift action of $\mathbb{Z}^2$ on $Y = \mathbb{Z}_2^{\mathbb{Z}^2}$ and $v = (1, 0)$. Then, $\Delta_v(\sigma)$ is the collection of all points x for which there exists a $k \in \mathbb{Z}$ with the property that $x(m, n) = 0$ whenever $m \geq k$. For explicit examples in a more general setting, see [2].

Proposition 6.8 *Let (X, σ) denote the 3-dot system. Then both $\Delta_{(-1,0)}(\sigma)$ and $\Delta_{(0,-1)}(\sigma)$ are infinite but $\Delta_{(-1,0)}(\sigma) \cap \Delta_{(0,-1)}(\sigma) = \{0\}$.*

Proof Let $\{a_i\}$ be an arbitrary sequence taking values in $\{0, 1\}$. From the defining property of the 3-dot system, it is easy to see that there exists a unique $x \in X$ such that $x(m, n) = 0$ whenever $m \geq 0$ and $x(-m, 0) = a_m$ for $m > 0$. Clearly any such x lies in $\Delta_{(-1,0)}(\sigma)$. Hence $\Delta_{(-1,0)}(\sigma)$ is infinite.

Similarly, there exists a unique $x \in X$ such that $x(m, n) = 0$ whenever $n \geq 0$ and $x(0, -n) = a_n$ for $n > 0$. This shows that $\Delta_{(0,-1)}(\sigma)$ is also infinite.

Now suppose x is an element of $\Delta_{(-1,0)}(\sigma) \cap \Delta_{(0,-1)}(\sigma)$. Since $x \in X$, we deduce that for all m, n and k, $x(m + 2^k, n) + x(m, n + 2^k) + x(m, n) = 0$. As the first two terms vanish for large values of k, we conclude that $x = 0$. □

We now show that the topological centraliser of the 3-dot system consists of algebraic maps. This is a form of topological rigidity. Similar rigidity properties holds even in the measure theoretic setting for a large class of actions of discrete groups [1, 3].

Theorem 6.9 *Let (X, σ) denote the* 3*-dot system and let $f : X \longrightarrow X$ be a continuous $\mathbb{Z}^2$-equivariant map. Then f is a continuous homomorphism.*

Proof We define $\widehat{f} : X \times X \longrightarrow X$ by $\widehat{f}(x, y) = f(x + y) - f(x) - f(y) + f(0)$. It is easy to see that $\widehat{f}$ is a $\mathbb{Z}^2$-equivariant map from $X \times X$ to X. Since $\widehat{f}$ is continuous and $X \times X$ is compact, it is also uniformly continuous.

It is easy to see that $\widehat{f}(x, y) = 0$ whenever $x = 0$ or $y = 0$. From uniform continuity of $\widehat{f}$, it follows that if $x \in \Delta_{(-1,0)}(\sigma)$ and $y \in \Delta_{(0,-1)}(\sigma)$ then $\widehat{f}(x, y)$ lies in $\Delta_{(-1,0)}(\sigma) \cap \Delta_{(0,-1)}(\sigma)$. As every infinite shift-invariant subgroup of X is dense, from the previous proposition, we deduce that $\Delta_{(-1,0)}(\sigma) \times \Delta_{(0,-1)}(\sigma)$ is a dense subgroup of $X \times X$, and $\widehat{f}$ maps it to $\{0\}$. Hence $\widehat{f}$ is identically zero.

This implies that f is affine, i.e., there exists a continuous homomorphism $\theta : X \longrightarrow X$ and $b \in X$ such that $f(x) = \theta(x) + b$. As $b = f(0)$ and f is shift equivariant, it follows that b is invariant under the shift action. Hence $b = 0$ and f is a continuous homomorphism. □

References

1. Bhattacharya, S. (2008). Isomorphism rigidity of commuting automorphisms. *Transactions of the American Mathematical Society, 360,* 6319–6329.
2. Einsiedler, M., Lind, D., Miles, R., & Ward, T. (2001). Expansive subdynamics for algebraic $\mathbb{Z}^d$-actions. *Ergodic Theory and Dynamical Systems, 21,* 1695–1729.
3. Lindenstrauss, E. (2005). Rigidity of multiparameter actions. *Israel Journal of Mathematics, 149,* 199–226.
4. Ledrappier, F. (1978). Un champ Markovien peut être d'entropie nulle et mélangeant. *Comptes Rendus de l'Académie des Sciences Paris Série A-B, 287*, 561–563.
5. Lind, D., & Marcus, B. (1995). *An introduction to symbolic dynamics and coding*. Cambridge University Press.
6. Lagarias, J., & Wang, Y. F. (1996). Tiling the line with translates of one tile. *Inventiones Mathematicae, 124*, 341–365.
7. Schmidt, K. (1995). *Dynamical systems of algebraic origin*. Birkhauser Verlag.
8. Szegedy, M. (1998). Algorithms to tile the infinite grid with finite clusters. *Proceedings of the 39th Annual Symposium on Foundations of Computer Science (FOCS '98)* (pp. 137–145).

Complex Dynamics

S. Sridharan and K. Verma

1 Introduction

These notes are based on a set of lectures given by the second author at the Advanced Instructional School on Ergodic Theory and Dynamical Systems held at IIT Delhi in December 2017. The goal of these lectures was to introduce the audience, that comprised mainly of PhD students, to some basic ideas in complex dynamics in one and several variables. No prior knowledge in dynamics was assumed, nor any originality in the presentation was claimed. The same applies to what follows. In fact, a good fraction of the course was based on the material in Beardon [2] and Steinmetz [9]. The last part on the dynamics of Hénon maps is a summary of some of the work begun in Bedford-Smillie [3]. Other aspects of the dynamics of this class of maps can be found in Fornaess-Sibony [4–6].

2 Some Preliminaries from Complex Analysis and Motivation

Let $\mathbb{P}^1 := \mathbb{C} \cup \{\infty\}$ denote the Riemann sphere that is defined as the complex plane along with the point at infinity. This is possible through the one point compactification of the stereographically projected plane onto the unit sphere $\mathbb{S}^2$ in $\mathbb{R}^3$. We denote the spherical metric on $\mathbb{P}^1$ as

S. Sridharan (✉)
Indian Institute of Science Education and Research Thiruvananthapuram IISER-TVM, Thiruvananthapuram, India
e-mail: shrihari@iisertvm.ac.in

K. Verma
Indian Institute of Science IISc, Bengaluru, India
e-mail: kverma@iisc.ac.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_5

$$\sigma(z, w) \;:=\; 2\frac{|z-w|}{\sqrt{1+|z|^2}\sqrt{1+|w|^2}}, \quad \text{for } z \neq \infty \text{ and } w \neq \infty.$$

Suppose one of the points, say $z = \infty$, one may consider the limit in the above definition, i.e., $\lim_{z\to\infty} \sigma(z, w)$. In these notes, we shall consider $\mathbb{P}^1$ to be our phase space, where we define functions, observe and understand the long term behaviour of its family of iterates.

By a rational map, we mean the map can be expressed as a quotient of two relatively prime polynomials, $R(z) := \dfrac{P(z)}{Q(z)}$. We define the degree of the rational map R, denoted as $\deg(R)$, to be the maximum among the degrees of the polynomials that yield the rational map; $\deg(R) := \max\{\deg(P), \deg(Q)\}$.

We briefly explain the root-finding algorithm for a real-valued polynomial, due to Isaac Newton and Joseph Raphson, known as the Newton-Raphson method in Numerical analysis. This provides the motivation for looking at the iterates of complex rational functions.

Let P be a real polynomial. Suppose, we start with $x = x_0$ as the initial guess for the solution of the equation $P(x) = 0$. Then, with appropriate conditions, the sequence $\{x_n\}_{n\geq 1}$ defined by

$$x_n \;:=\; x_{n-1} - \frac{P(x_{n-1})}{P'(x_{n-1})},$$

converges to a solution of $P(x) = 0$. Observe that defining

$$f(x) \;:=\; x - \frac{P(x)}{P'(x)},$$

we have, in the Newton-Raphson method,

$$x_n \;=\; f^n(x_0), \;\; \text{where } f^n \;:=\; \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times}}.$$

Thus, obtaining a root of the polynomial P with x_0 being the initial guess is equivalent to studying the long-term behaviour of the orbit of the point x_0 under the iterates of the function f, i.e., $\{f^n(x_0) : n \geq 0\}$. Further, finding a root of the polynomial P is equivalent to finding a fixed point for the function f.

Cayley proposed to apply this method to a complex polynomial $P(z)$. For example, consider $P(z) = z^2 - 1$. Then,

$$f(z) \;=\; z - \frac{z^2-1}{2z} \;=\; \frac{z^2+1}{2z}.$$

The fixed points of f are the solutions of the equation $f(z) = z$; in this case $z = \pm 1$. We will now investigate the behaviour of the function f locally near $z = 1$, one of its fixed points and globally.

Consider the Taylor series of f about $z = 1$ given by

$$f(z) \;\backsimeq\; f(1) + \frac{f'(1)(z-1)}{1!} + \frac{f''(1)(z-1)^2}{2!} + \cdots .$$

Let $\varphi(z) = z + 1$ and $\tilde{f} = \varphi^{-1} \circ f \circ \varphi$. Then, $\tilde{f}$ has a fixed point at $z = 0$ satisfying $\tilde{f}'(0) = 0$. Hence, the Taylor series of $\tilde{f}$ about $z = 0$ is given by

$$\tilde{f}(z) \;=\; a_2 z^2 + a_3 z^3 + \cdots .$$

Further, making a change of variables by defining $\psi(z) := \beta z$, where $\beta \neq 0$ and putting $g(z) := \psi^{-1} \circ \tilde{f} \circ \psi(z)$, we have

$$g(z) \;=\; a_2 \beta z^2 + \cdots .$$

Making an appropriate choice for $\beta \neq 0$, we ensure that $|g(z)| < |z|^2$ in a sufficiently small neighbourhood around $z = 0$, in order that we obtain, by induction, that

$$\left|g^n(z)\right| \;<\; |z|^{2^n} \quad \text{for} \quad n \geq 0.$$

It is then quite clear that for $r < 1$, the iterates of any point in the disc $D(0, r)$ uniformly converges to 0. Tracing back the change of variables, we see that the sequence, $\{f^n(z)\} \to 1$ uniformly. An analogous analysis is also true if one investigates the behaviour of points near the other fixed point -1. The behaviour of the iterates of f seem to be stable locally near the fixed points of the function.

We will now study the behaviour of f, globally. We write

$$f(z) \;=\; \frac{z^2+1}{2z} \;=\; \frac{1}{2}\left(z + \frac{1}{z}\right).$$

Consider the Möbius transformation $\phi(z) := \dfrac{z-1}{z+1}$. Then ϕ conjugates f to the the map $z \longmapsto z^2$ in the sense that,

$$\phi^{-1} \circ f \circ \phi(z) \;=\; z^2.$$

$$\begin{array}{ccc} \mathbb{P}^1 & \xrightarrow{\quad f \quad} & \mathbb{P}^1 \\ \uparrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \phi} \\ \mathbb{P}^1 & \xrightarrow{\quad z \mapsto z^2 \quad} & \mathbb{P}^1 \end{array}$$

It follows that

$$\phi^{-1} \circ f^n \circ \phi(z) \;=\; z^{2^n}.$$

Thus, in order to study the iterates of f, we consider the iterates of the map $h(z) = z^2$, i.e., $\{h^n(z)\}_{n\geq 0}$. Observe that $\{h^n(z) = z^{2^n}\}$ uniformly converges to the constant function 0 in the spherical metric, on the open unit disc $|z| < 1$ and to the constant map ∞ in the same metric, on the exterior of the closed unit disc $|z| > 1$. Moreover, it is easily verifiable that ϕ maps the unit circle in $\mathbb{C}$ onto the imaginary axis, the open unit disc onto the left half-plane and the exterior of the closed unit disc onto the right half-plane.

Since $0,\ \infty$ are the fixed points of the map h, it follows that $\phi(0) = -1$ and $\phi(\infty) = 1$. Hence, we have

$$f^n(z) = \phi^{-1} \circ h^n \circ \phi(z) \longrightarrow -1, \quad \text{if } \operatorname{Re}(z) < 1;$$
$$f^n(z) = \phi^{-1} \circ h^n \circ \phi(z) \longrightarrow \;\; 1, \quad \text{if } \operatorname{Re}(z) > 1.$$

Question: What happens to the iterates of $f(z)$ on the imaginary axis or equivalently what happens to the iterates of $h(z)$ on the unit circle?

For any point $z = e^{i\theta} \in \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$, $h(z) = e^{2i\theta}$ and so $h^n(z) = e^{2^n i\theta}$. Suppose $\theta = 2\pi p/2^m$ for some $p, m \in \mathbb{Z}_+$, then $h^m(z) = 1$. Since 1 is a fixed point of h, we also have $h^{m+k}(z) = 1$ for every $k \geq 0$. These points are the 2^mth roots of unity in the complex plane. Note that the set $\{z \in \mathbb{S}^1 : z$ is a 2^mth root of unity, $m \geq 1\}$ forms a dense subset of unit circle $\mathbb{S}^1$. However, if θ is an irrational multiple of 2π, then the set $\{h^n(z) : n \geq 1\}$ is dense in $\mathbb{S}^1$. It is clear that the behaviour of the iterates of $h(z) = z^2$ on the unit circle $\mathbb{S}^1$ seems to be quite complicated.

3 Normal Families and Dichotomy of $\mathbb{P}^1$

We begin this section with the definition of normal families.

Definition 3.1 Let $\Omega \subseteq \mathbb{P}^1$ be a domain and $\mathcal{F}$, a family of continuous functions defined on Ω. We say that $\mathcal{F}$ is a *normal family* if every infinite sequence in $\mathcal{F}$ has a subsequence that converges uniformly on all compact subsets of Ω.

Theorem 3.2 (Arzela-Ascoli Theorem) *A family $\mathcal{F}$ of continuous functions defined on a region $\Omega \subset \mathbb{P}^1$ that take values in $\mathbb{P}^1$ is normal iff $\mathcal{F}$ is equicontinuous on every compact subset $K \subset \Omega$.*

Thus, in this case, equicontinuity and normality are equivalent conditions—see [2]. The point to be understood is that normality is an analogue of compactness. We have the following theorems, due to Montel, who initiated the study of normal families.

Theorem 3.3 (Montel's Theorem I) *A family $\mathcal{F}$ of holomorphic functions defined on a domain $\Omega \subseteq \mathbb{P}^1$ is normal if every $f \in \mathcal{F}$ is uniformly bounded, i.e., there exists $M > 0$ such that $|f(z)| \leq M$ for all $z \in \Omega$ and $f \in \mathcal{F}$.*

Proof This is a consequence of the Cauchy Integral Formula. Let z_0, w_0 be sufficiently close to each other. Let D be some open disc such that $\overline{D} \subset \Omega$, containing the points z_0, w_0. Then,

$$f(z_0) = \frac{1}{2\pi i}\int_{\partial D}\frac{f(z)}{z-z_0}dz, \qquad f(w_0) = \frac{1}{2\pi i}\int_{\partial D}\frac{f(z)}{z-w_0}dz.$$

So,

$$\begin{aligned}|f(z_0)-f(w_0)| &\leq \frac{1}{2\pi}\int_{\partial D}\frac{|f(z)||z_0-w_0|}{|(z-w_0)(z-z_0)|}|dz| \\ &\leq \frac{M}{2\pi}\int_{\partial D}\frac{|z_0-w_0|}{|(z-w_0)(z-z_0)|}|dz|.\end{aligned}$$

Hence, for any $\epsilon > 0$, we can choose $\delta > 0$ such that $|f(z_0) - f(w_0)| \leq \epsilon$, whenever $|z_0 - w_0| < \delta$ for all $f \in \mathcal{F}$, proving $\mathcal{F}$ is equicontinuous and thereby, normal. □

We now define the term covering space of a topological space, that we will use in the proof of the next version of Montel's theorem.

Definition 3.4 A *covering space* of a topological space X is a topological space Y together with a continuous surjective map $\pi : Y \longrightarrow X$ such that for every $x \in X$ there exists an open neighbourhood U of x satisfying the condition that $\pi^{-1}(U)$ can be expressed as a union of disjoint open sets in Y, each of which is mapped homeomorphically onto U by π.

Here are two very useful versions of Montel's theorem for meromorphic functions. Proofs and further variants can be found in [2].

Theorem 3.5 (Montel's Theorem II) *Let $\mathcal{F}$ be a family of meromorphic functions defined on a domain Ω that omits 0, 1 and ∞. Then $\mathcal{F}$ is normal.*

Proof Instead of dealing with $f : \Omega \longrightarrow \mathbb{P}^1 \setminus \{0, 1, \infty\}$, we study $f : \Omega \longrightarrow \mathbb{C}\setminus\{0, 1\}$. The covering space of $\mathbb{C}\setminus\{0, 1\}$ is the unit disc $\mathbb{D}$. Every $f \in \mathcal{F}$ admits a lift locally, say on a disc in Ω and if $\tilde{f}$ is this lift, then the family of lifts forms a normal family since they take values in $\mathbb{D}$. Since $\pi \circ \tilde{f} = f$, it follows that $\mathcal{F}$ is a normal family. □

Theorem 3.6 (Montel's theorem III) *Let $\mathcal{F}$ be a family of holomorphic functions on a domain $\Omega \subset \mathbb{P}^1$ such that every $f \in \mathcal{F}$ omits a set of three distinct points $\{a_f, b_f, c_f\}$. If the spherical distances between the pairs (a_f, b_f), (b_f, c_f) and (c_f, a_f), as f varies in $\mathcal{F}$, are uniformly bounded below by a positive constant, then $\mathcal{F}$ is normal on Ω.*

Definition 3.7 A point $z_0 \in \mathbb{P}^1$ is a *branch point* of a rational map R if for every neighbourhood around z_0, R is not a homeomorphism restricted to the neighbourhood.

Let R be a rational map with degree d. Then the number of branch points of R, counting multiplicity is equal to $2(d-1)$. A rational map $R : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ defines a $d:1$ branched covering. In what follows, we use the notation $R^{(k)}$ and R^n to denote the kth differential and the nth iterate of the rational map R respectively.

Assume without loss of generality that $z_0 = 0$ is a branch point of R that satisfies $R(0) = 0$. Since the branch points are also the critical points of R, we have $R(0) = 0,\ R'(0) = 0,\ \ldots,\ R^{(k-1)}(0) = 0$ whereas $R^{(k)}(0) \neq 0$, for some $k \geq 2$. Thus, in a neighbourhood around the point $z = 0$, we have

$$R(z) \ = \ a_k z^k + \cdots \ = \ z^k h(z),$$

where $h(z)$ is holomorphic at $z = 0$ and $h'(0) \neq 0$. Thus, $h(z) = (g(z))^k$, for some holomorphic function $g(z)$. Hence, in the considered neighbourhood around $z = 0$, we have $R(z) = (zg(z))^k$. In the new coordinate system $w = zg(z)$, this equation becomes $R(w) = w^k$. Thus R is of degree k, locally near its branch points.

We shall henceforth focus on the global behaviour of the rational map R on $\mathbb{P}^1$. We start with the definition of the Fatou and the Julia set that dichotomises $\mathbb{P}^1$ based on the equicontinuity of the family of iterates of the considered rational map R.

Definition 3.8 Let $R : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ be a rational map. Then the largest open set of $\mathbb{P}^1$, where the family of iterates of R, namely $\{R^n : n \geq 1\}$ is equicontinuous is called the *Fatou set* of R and is denoted by F_R. Its complement $\mathbb{P}^1 \setminus F_R$ is called the *Julia set* and is denoted by J_R.

Observe that, by definition, J_R is a closed subset of the compact set $\mathbb{P}^1$ and hence, compact while F_R is an open set. We urge the reader to observe that we can alternately define the Fatou and the Julia set in a more convenient way, by the concept of normality instead of equicontinuity using the Arzela-Ascoli theorem.

We now enlist a few basic properties of the Fatou set and the Julia set.

Theorem 3.9 *Let R be a rational map with* $\deg(R) \geq 2$. *Then $J_R \neq \varnothing$.*

Proof We prove this by contradiction. Suppose $J_R = \varnothing$, then $F_R = \mathbb{P}^1$. This implies that the family $\{R^n : n \geq 1\}$ is normal on the entire Riemann sphere, $\mathbb{P}^1$. Then the family should converge to a meromorphic limit function, say S, at least for some subsequence $\{R^{n_k} : k \geq 1\}$ of $\{R^n : n \geq 1\}$. The function S, being meromorphic in $\mathbb{P}^1$, is nothing but a rational map and therefore must be of some finite degree, say d'. However, $\deg(R^{n_k}) \nrightarrow \deg(S)$. In fact, $\deg(R^{n_k})$ grows exponentially as $k \to \infty$, since $\deg(R) \geq 2$. This contradiction proves that the Julia set can not be empty; $J_R \neq \varnothing$. □

Upon proving that the Julia set is never empty for a rational map R of degree ≥ 2, one natural question that arises is whether there exists rational maps with an empty

Fatou set. The answer to this question is in the affirmative. An example of such a map was constructed by a French mathematician, Samuel Lattès in 1918, namely

$$R(z) := \frac{(z^2+1)^2}{4z(z^2-1)}.$$

We will study further about this example, in a later section.

Definition 3.10 Let R be a rational map. A domain Ω is said to be

- *forward invariant* under the map R, if $R(\Omega) \subset \Omega$;
- *backward invariant* under the map R, if $R^{-1}(\Omega) \subset \Omega$; and
- *completely invariant* under the map R if it is both forward and backward invariant under the map R.

Note that Ω is completely invariant under R if and only if Ω^c is completely invariant under R.

Theorem 3.11 *Let R be a rational map of degree at least 2. Then the Fatou set F_R and the Julia set J_R are completely invariant under R.*

Proof Here, we will only prove that F_R is forward invariant. The proof of backward invariance of F_R is analogous. Further owing to the remark before the start of this theorem, we then know that J_R is completely invariant too.

Let $p \in F_R$. Then, there exists a neighbourhood $D(p, r) \subset F_R$, where the family $\{R^n : n \geq 1\}$ is normal. Since R is an open map, $R(D(p, r))$ is an open neighbourhood of $R(p)$. Further, the family $\{R^n : n \geq 1\}$ is normal in a neighbourhood of p and hence, the family $\{R^{n-1} : n \geq 1\}$ is normal in a neighbourhood of $R(p)$. Thus, $R(p) \in F_R$. □

Theorem 3.12 *Fix $k > 0$. Then $F_{R^k} = F_R$ and $J_{R^k} = J_R$.*

Proof Observe that the family of iterates of R^k, i.e., $\{R^{kn} : n \geq 1\}$ is a subfamily of the family of iterates of R, i.e., $\mathcal{F} = \{R^n : n \geq 1\}$. Hence $F_R \subseteq F_{R^k}$.

To obtain the other way inclusion, we first observe that R^m is uniformly continuous for every m, in the spherical metric. Thus, for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sigma(R^m(x), R^m(y)) < \epsilon \quad \text{whenever} \quad \sigma(x, y) < \delta. \tag{1}$$

The family $\{R^{kn} : n \geq 1\}$ is equicontinuous on F_{R^k}. This implies that for every $\delta > 0$ there exists a $\delta_1 > 0$ such that

$$\sigma(R^{nk}(x), R^{nk}(y)) < \delta \quad \text{whenever} \quad \sigma(x, y) < \delta_1 \ \ \forall\, n \geq 1. \tag{2}$$

By equation (1), we have that

$$\sigma(R^m \circ R^{nk}(x), R^m \circ R^{nk}(y)) < \epsilon \quad \text{whenever} \quad \sigma(R^{nk}(x), R^{nk}(y)) < \delta.$$

By equation (2), the family $\mathcal{F}_m = \{R^m R^n : n \geq 0\}$ is equicontinuous on the Fatou set F_{R^k} for every integer $m \geq 0$. Hence the finite union $\mathcal{F}_0 \cup \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_{k-1}$ is also equicontinuous on the Fatou set F_{R^k}. However, $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_{k-1}$. □

4 Rational Maps with Empty Fatou Set

We begin this section with the definition of a complex period for a meromorphic function.

Definition 4.1 Let f be a meromorphic function on the entire plane. A non-zero complex number w is said to be a *period* for f if $f(z + w) = f(z)$ for every $z \in \mathbb{C}$.

Note that if w is a period, then so are its multiples nw where n is an integer. The set Λ of all the periods of a given f then clearly forms a module over the integers. $\Lambda \subset \mathbb{C}$ is also discrete as otherwise the identity theorem and the fact that $f(w) = f(0)$ for all $w \in \Lambda$ will imply that f is constant.

It is possible to describe the structure of Λ. Indeed, Theorem (1) of Chap. 7 in Ahlfors [1] shows that if Λ contains a non-zero element, then every element in it can be written as nw, where $w \neq 0$ and n is an integer or as $n_1 w_1 + n_2 w_2$, where w_1, w_2 are a pair of non-zero complex numbers with w_2/w_1 non-real and n_1, n_2 are integers. In the latter case,

$$\mathfrak{F} = \{s w_1 + t w_2 : 0 \leq s, t < 1\}$$

is called the *period (or fundamental) parallelogram* and Λ will be referred to as a *lattice*.

For example, the bounded region in the picture below corresponds to a fundamental parallelogram for an appropriately defined Λ (Fig. 1).

It should be noted that the set of periods Λ corresponding to an entire function can never admit a fundamental parallelogram. If it does, it necessarily has to be a constant.

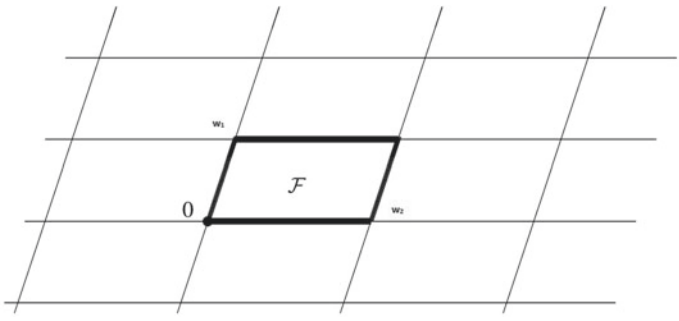


Fig. 1 Fundamental mesh

To see this, suppose f is an entire function with a fundamental parallelogram $\mathfrak{F}$. By definition, it follows that $f(\mathbb{C}) = f(\mathfrak{F})$ and since the latter is bounded, Liouville's theorem shows that f is necessarily constant. However, every element in the set of periods Λ for $f(z) = \sin z$ for example is an integral multiple of 2π.

Since this argument clears the non-existence of a non-constant entire function with two fundamental periods, the next natural question is to investigate the existence of a meromorphic function with two fundamental periods. Weierstrass constructed a doubly periodic meromorphic function on $\mathbb{C}$. Let us recall a basic construction. For example, see Ahlfors [1] for details.

Lemma 4.2 *Let $\Lambda \subset \mathbb{C}$ be a lattice. The series*

$$\sum_{\lambda \in \Lambda \setminus \{0\}} |\lambda|^{-3}$$

converges.

This implies that

Theorem 4.3 *The series*

$$\wp(z) \;:=\; \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}$$

converges locally uniformly absolutely in $\mathbb{C} \setminus \Lambda$. It defines a meromorphic Λ-periodic function, called the Weierstrass' $\wp$-function.

Use the above lemma to obtain $\wp$ as a meromorphic function on $\mathbb{C}$. Then, the poles of $\wp$ are precisely at the lattice points.

Definition 4.4 A function f is said to be an *elliptic function* if it satisfies the following conditions:

1. f is doubly periodic with respect to some lattice Λ;
2. f is meromorphic on $\mathbb{C}$.

Let $\langle w_1, w_2 \rangle_{\mathbb{Z}} = \Lambda$ denote the group generated by $\{w_1, w_2\}$ over $\mathbb{Z}$. Identify the opposite sides of the fundamental mesh of Λ to obtain a torus $\mathbb{T}$, which is a compact Riemann surface, i.e., $\mathbb{C}/\langle w_1, w_2 \rangle_{\mathbb{Z}} \simeq \mathbb{T}$. Thus $\wp$ is a mapping from $\mathbb{T}$ to $\mathbb{P}^1$. As $\mathbb{T}$ is compact, $\wp(\mathbb{T})$ is closed. Also $\wp(\mathbb{T})$ is open (by the Open Mapping Theorem). As $\mathbb{P}^1$ is connected, $\wp(\mathbb{T}) = \mathbb{P}^1$ is surjective (Fig. 2).

Theorem 4.5 *There exists a rational map $R : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ such that*

$$\wp(2z) = R \circ \wp(z); \tag{3}$$

$$R(z) = \frac{z^4 + g_2\left(\frac{z^2}{2}\right) + 2g_3 z + \left(\frac{g_2}{4}\right)^2}{4z^3 - g_2 z - g_3}, \tag{4}$$

where g_2, g_3 are constant terms depending on the lattice.

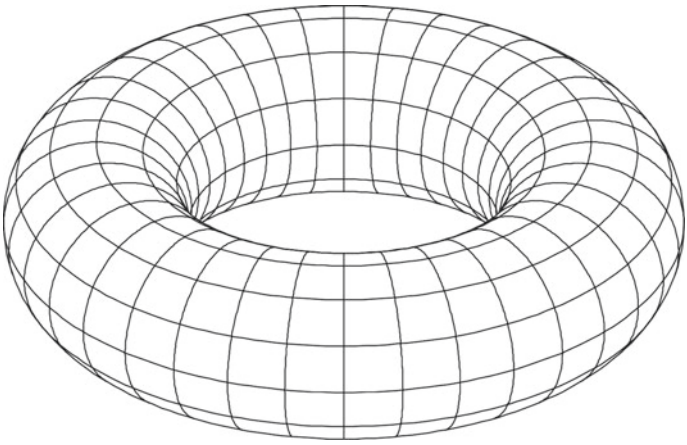


Fig. 2 Torus

We sketch some of the details concerning the function $\wp$ and justify the formula in (3). Appealing to lemma (4.2), one can see that the function $\wp$ is meromorphic on $\mathbb{C}$. As $\mathbb{T}$ and $\mathbb{P}^1$ are compact Riemann surfaces, $\wp$ acts as a branched covering map. By considering the poles of $\wp$, we see that $\wp$ is a 2-sheet covering i.e., for each $w \in \mathbb{P}^1$, there are exactly two solutions (modulo the lattice Λ) of $\wp(z) = w$ in $\mathbb{C}$. Given any $w \in \mathbb{P}^1$, there are only two solutions z_1, z_2 of the equation $\wp(z) = w$. These can be taken to be, say, u and $w_1 + w_2 - u$. Then,

$$\wp(2[(w_1 + w_2) - u]) \;=\; \wp(2u). \tag{5}$$

It is then clear that $\wp$ is an even function, that is $\wp(z) = \wp(-z)$. Further, owing to (5), one might expect a formula of the type $\wp(2z) = R(\wp(z))$ to hold. Referring to (5), we can define the map $w \longmapsto \wp(2u)$ of $\mathbb{P}^1$ onto itself, independent of the choice of u. It is then easy to see that this map is meromorphic, thus must be a rational map, say R. This gives us a motivation to expect a formula, $R(\wp(z)) = \wp(2z)$.

Proof We will now sketch the details of the proof of the formula in (4). It is very clear that the derivative of $\wp$, namely $\wp'$ has triple poles only at the lattice points in Λ. Because of this, it is not hard to see that one can construct a cubic polynomial P such that the elliptic function $\wp'(z)^2 - P(\wp(z))$ has no poles at the origin. Hence, there exists no poles for $\wp'(z)^2 - P(\wp(z))$ in $\mathbb{C}$, making it bounded. A computation of P then leads to the relation

$$\wp'(z)^2 \;=\; 4\wp(z)^3 - g_2\wp(z) - g_3, \tag{6}$$

where g_2, g_3 are constants depending on the lattice Λ and are given by

$$g_2 \;=\; 60 \sum_{\lambda\in\Lambda\setminus\{0\}} \frac{1}{(\lambda)^4} \quad \text{and} \quad g_3 \;=\; 140 \sum_{\lambda\in\Lambda\setminus\{0\}} \frac{1}{(\lambda)^6}.$$

Now, select distinct points u and v in $\mathbb{C}$ where $\wp$ takes different values. We determine the values of A and B satisfying

$$\wp'(u) \;=\; A\wp(u) + B, \quad \text{and} \quad \wp'(v) \;=\; A\wp(v) + B,$$

to be

$$A \;=\; \frac{\wp'(u) - \wp'(v)}{\wp(u) - \wp(v)} \quad \text{and} \quad B \;=\; \frac{\wp(v)\wp'(u) - \wp(u)\wp'(v)}{\wp(v) - \wp(u)}.$$

It is then clear that $f(z) = \wp'(z) - A\wp(z) - B$ has three poles in Λ. And consequently, f must have three zeroes (by the argument principle). By construction, two of these zeroes occur at u and v. A consequence of the argument principle states that $\sum p_i$ differs from $\sum z_i$ by an element of Λ, where p_i's are the poles of f and z_i's are the zeroes of f. In our case, all the poles of f occur at the origin. This implies $u,\ v$ and $-(u+v)$ are the zeroes of f. However, since

$$[f(z) + A\wp(z) + B]^2 \;=\; \wp'(z)^2 \;=\; 4\wp(z)^3 - g_2\wp(z) - g_3,$$

we find that $\wp(u),\ \wp(v)$ and $\wp(-(u+v))$ are the solutions of the equation

$$[Az + B]^2 \;=\; 4z^3 - g_2 z - g_3.$$

Hence,

$$\wp(u) + \wp(v) + \wp(-(u+v)) \;=\; \frac{A^2}{4} \;=\; \frac{1}{4}\left(\frac{\wp'(u) - \wp'(v)}{\wp(u) - \wp(v)}\right)^2.$$

Now, letting $u \to v$ and using the fact that the function $\wp$ is even, we obtain

$$2\wp(v) + \wp(2v) \;=\; \frac{A^2}{4} \;=\; \frac{1}{4}\left(\frac{\wp''(v)}{\wp'(v)}\right)^2. \tag{7}$$

Finally, differentiating both sides of (6) gives an expression for $\wp''(z)$. Using this expression together with (6) and (7), we obtain the addition formula given by (3) and (4). □

We wish to include a different argument sketched by the referee that shows the existence of such a rational function. It does not quite give a formula for it but it is of independent interest since it is based on a pole-counting argument that is ubiquitous in the study of such functions.

The function $\wp(2z)$ is an even elliptic function with respect to the lattice Λ and has double poles at the points of $(1/2)\Lambda$. Let Λ be generated by w_1, w_2 and consider

$$g(z) = (\wp(z) - \wp(w_1/2))^2 \cdot (\wp(z) - \wp(w_2/2))^2 \cdot (\wp(z) - \wp((w_1 + w_2)/2))^2 \cdot \wp(2z).$$

This is an even elliptic function and since the poles of $\wp(2z)$ are cancelled by the zeros of the other three factors, the only pole in the fundamental parallelogram is at the origin and it is of order 14 (each factor contributes an order of 4). Thus, near the origin, the singular part of g looks like

$$\frac{a_7}{z^{14}} + \frac{a_6}{z^{12}} + \cdots + a_0.$$

Each of these can be cancelled out recursively as follows. The function $g_1 = g - a_7(\wp)^7$ has a pole of order 12 at the origin and hence its singular part looks like

$$\frac{b_6}{z^{12}} + \cdots + b_0.$$

Repeating this process by subtracting powers of $\wp(z)$ leads to an elliptic function that has no poles at the origin and hence it must constant. This shows that there is a polynomial P such that

$$g \;=\; P(\wp) + C$$

for some constant C and this when rearranged shows that $\wp(2z)$ can be expressed as a rational function of $\wp(z)$.

Thus, the theorem asserts the existence of a holomorphic map $R : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ such that the following diagram commutes.

$$\begin{array}{ccc} \mathbb{T} & \xrightarrow{z \mapsto 2z} & \mathbb{T} \\ \Big\downarrow{\scriptstyle \wp} & & \Big\downarrow{\scriptstyle \wp} \\ \mathbb{P}^1 & \xrightarrow{R} & \mathbb{P}^1 \end{array}$$

Samuel Lattés proved, in 1918, that the Fatou set of the above map R is an empty set. We will show that the family $\{R^n\}_{n\geq 1}$ is not normal in any neighbourhood of any point in $\mathbb{P}^1$.

Consider an open set D in $\mathbb{P}^1$. Then, $U = \wp^{-1}(D)$ is an open set. Applying the doubling map $z \longmapsto 2z$, n times on U, we note that it gets expanded by a factor of 2^n. Thus, for sufficiently large n, the image of U under the doubling map contains the fundamental mesh, i.e.,

$$R^n(D) \;=\; R^n(\wp(U)) \;=\; \wp(2^n U) \;=\; \mathbb{P}^1.$$

This implies that R^n expands any small open set D onto $\mathbb{P}^1$. Hence, the family $\{R^n\}_{n\geq 1}$ cannot be normal in any open set of $\mathbb{P}^1$ implying $J_R = \mathbb{P}^1$. Indeed, $R^n(D) = \mathbb{P}^1$ implies that $R^n(D)$ must intersect the Julia set J_R. The backward invariance of J_R

shows that D must also intersect it (the Julia set) and since it is closed and D is arbitrary, it must be the case that $J_R = \mathbb{P}^1$.

Note that for a suitable choice of the lattice Λ, we obtain $g_2 = 4$ and $g_3 = 0$. Then, the Lattès map is given by

$$R(z) \;=\; \frac{(z^2+1)^2}{4z(z^2-1)}.$$

5 Some Properties of the Julia Set

In this section, we shall investigate some elementary properties of the Julia set of a rational map, R.

Definition 5.1 For $z_0 \in \mathbb{P}^1$, consider the Taylor series of R about z_0, i.e.,

$$R(z) \;=\; R(z_0) + \sum_{j \geq 1} a_j \, (z - z_0)^j.$$

Then the *branching order (or valency)* of R at z_0 is the minimum j for which $a_j \neq 0$, denoted by $\nu(z_0)$.

For most points in $\mathbb{P}^1$, we have $\nu(z) = 1$. This is not true only for finitely many points in $\mathbb{P}^1$. Here is a result that captures this idea.

Theorem 5.2 (Riemann-Hurwitz Formula) *[9] Let R be a rational map with* $\deg(R) = d$. *Then*

$$\sum_{z \in \mathbb{P}^1} (\nu(z) - 1) \;=\; 2d - 2.$$

Let $R : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ be a rational map, as earlier of degree $d \geq 2$. Define a relation $\sim$ on $\mathbb{P}^1$ as follows: For $x, y \in \mathbb{P}^1$ we define $x \sim y \iff R^n(x) = R^m(y)$, for some positive integers m and n. It is a simple exercise to note that $\sim$ is an equivalence relation on $\mathbb{P}^1$.

Let $[x]$ denote the equivalence class of $x \in \mathbb{P}^1$. Since $R^n(x) = R^m(y) \rightarrow R^{n+1}(x) = R^{m+1}(y)$ and R is surjective, we see that $[x]$ is completely invariant.

Proposition 5.3 *If V is a finite, completely invariant set under a rational map R, then V contains at most two points.*

Proof Let V be a finite, completely invariant set under the rational map R with cardinality k. Our aim here is to prove that $k \leq 2$. Since V is completely invariant under the action of the rational map R, we have $R(V) = V = R^{-1}(V)$. Then R acts as permutation on V and therefore, for some $m \geq 1$, we have $R^m : V \longrightarrow V$ to be identically equal to the identity permutation, i.e., R^m fixes each point of V. We know by the Riemann-Hurwitz Formula,

$$\sum_{z \in V} (\nu(z) - 1) \;\leq\; \sum_{z \in \mathbb{P}^1} (\nu(z) - 1) \;=\; 2d^m - 2,$$

where d is the degree of $R(z)$. Since, for each $z \in V$, $R^m(z) = z$, $(R^m)^{-1}(z) = \{z\}$ and the degree of R^m is equal to d^m, we have $\nu(z) = d^m$. This implies

$$k(d^m - 1) \;\leq\; 2(d^m - 1) \to k \;\leq\; 2.$$

□

Definition 5.4 We say that $x_0 \in \mathbb{P}^1$ is an *exceptional point* of R, if $[x_0]$ is finite. Let $E := \{x : x \text{ is an exceptional point}\}$. Then, E is said to be an *exceptional set* of R.

An immediate observation from proposition (5.3) implies that the cardinality of $[x]$ i.e., $\#([x]) \leq 2$, whenever x is an exceptional point.

Theorem 5.5 *A rational map R of degree at least 2 has atmost 2 exceptional points. If E is a singleton, then R is conjugate (via a Möbius map) to a polynomial. If E contains two elements, then R is conjugate (via a Möbius map) to the map $z \longmapsto az^d$ where $a \in \mathbb{C}$. Furthermore, the exceptional set E of the rational map R is contained in its Fatou set F_R.*

Proof Since the exceptional set is completely invariant by definition, it is clear that E can contain at most two points, by proposition (5.3). Thus, there are four possibilities for E, namely:

1. $E = \varnothing$;
2. $E = \{\zeta_0\} = [\zeta_0]$;
3. $E = \{\zeta_1, \zeta_2\},\ [\zeta_1] = \{\zeta_1\},\ [\zeta_2] = \{\zeta_2\}$;
4. $E = \{\zeta_1, \zeta_2\},\ [\zeta_1] = [\zeta_2]$.

1. This is a trivial case, about which we do not say anything.
2. This means $R^{-1}(\zeta_0) = \{\zeta_0\}$. However, we know that the property $P^{-1}(\infty) = \{\infty\}$ characterises the polynomials among rational functions, since then P has exactly one pole at ∞ and no poles in the finite complex plane. Hence, R is conjugate to some polynomial P (via a Möbius map).
3. This means $R^{-1}(\zeta_1) = \{\zeta_1\}$ and $R^{-1}(\zeta_2) = \{\zeta_2\}$. We note again that for polynomials P of the from az^d, we have that $P^{-1}(0) = \{0\}$ and $P^{-1}(\infty) = \{\infty\}$. Hence, R is again conjugate (via a Möbius map) to some polynomial $z \longmapsto az^d$, where $a \in \mathbb{C}$ and d is some positive integer.
4. Here, we have $R(\zeta_1) = \zeta_2$ and $R(\zeta_2) = \zeta_1$. Consider the rational map $z \longmapsto az^d$, where $a \in \mathbb{C}$ and d is a negative integer, that interchanges the points at 0 and ∞. Hence, R is then conjugate (via a Möbius map) to the map $z \longmapsto az^d$, where $a \in \mathbb{C}$ and d is some negative integer.

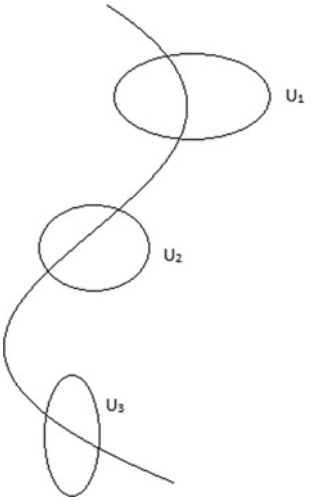
However, from all the above cases, it is clear that $E \subset F_R$. □

Theorem 5.6 *For a rational map R of degree at least 2, J_R is infinite.*

Proof Theorem (3.9) shows that $J_R \neq \varnothing$. Suppose $\xi \in J_R$. If J_R were finite, then ξ must be an exceptional point, since J_R is completely invariant. However, this is not possible because the exceptional set lies inside the Fatou set, which is the complex complement of J_R. Thus J_R should be infinite. □

Theorem 5.7 *Let R be a rational map of degree at least 2 and U be an open set such that $U \cap J_R$ is non-empty. Then there exists some $N > 0$ such that $R^N(U) \supseteq J_R$.*

Proof Let U_1, U_2 and U_3 be disjoint open sets in $\mathbb{P}^1$ that has a non-empty intersection with J_R.



Claim: For each $i \in \{1, 2, 3\}$, there exists some $N = N(i) \in \mathbb{Z}_+$ and $j = j(i) \in \{1, 2, 3\}$ such that $R^N(U_i) \supset U_j$.

We prove this claim by the method of contradiction. Suppose, this is not true, then the family of sets obtained by the action of the iterates of R on U_i, i.e., $R^n(U_i)$, do not cover either of the U_j's $i \neq j$ with $i, j \in \{1, 2, 3\}$. Without loss of generality, we can take U_1 to satisfy this condition. Then, the family $\{R^n|_{U_1} : n \geq 1\}$ leaves sufficiently more than three distinct points from its image set. Hence, by Montel's theorem II, this family is normal on U_1. This is a contradiction, since $U_1 \cap J_R \neq \varnothing$.

Since the claim is true, we can now choose N that satisfies $U_1 \subset R^N(U_1) \subset R^{2N}(U_1) \subset \cdots$. This is an increasing sequence of sets that acts as an open cover of J_R. Since $\{R^N, R^{2N}, \ldots\}$ is not normal in the whole of U_1, we expect it to leave atmost two points only (Montel's theorem II). However, J_R is compact, and hence, $J_R \subset R^m(U_1)$ for some $m \geq 1$. □

Theorem 5.8 $J_R \subset \overline{\{\textit{All periodic points of } R\}}$ *for a rational map R of degree at least 2.*

Proof Choose $w_0 \in J_R$ such that w_0 is not a critical point of R^2. Then there exists at least three distinct pre-images of w_0 under R^2, say w_1, w_2 and w_3. Choose mutually

disjoint neighbourhoods $U_0,\ U_1,\ U_2$ and U_3 of $w_0,\ w_1,\ w_2$ and w_3 respectively, such that $R^2\big|_{U_j} \longrightarrow U_0$ is a homeomorphism. Let

$$S_j(z) \ := \ \left(R^2\big|_{U_j}\right)^{-1}(z).$$

To complete the proof of the theorem, we now show that for some $z_0 \in U_0$, there exists an $m \geq 1$ such that $R^m(z_0) = S_j(z_0)$ for all $j \in \{1, 2, 3\}$. Suppose not, then $R^n|_{U_0}$ would be normal by Montel's theorem, which is not true. Hence,

$$R^m(z_0) \ = \ S_j(z_0) \quad \Longrightarrow \quad R^{m+2}(z_0) \ = \ z_0.$$

Thus, z_0 is a periodic point of R. Hence, every open neighbourhood U_0 that has a non-empty intersection with J_R contains a periodic point of R. □

6 Local Analysis Near a Fixed Point

Consider $\Omega \subset \mathbb{P}^1$ and $f : \Omega \longrightarrow \Omega$ be a holomorphic map. Suppose $z_0 \in \Omega$ is a fixed point of f. In this section, we investigate the local behaviour of f near its fixed point z_0.

Recall that the Taylor series expansion of f about the fixed point z_0 is given by,

$$f(z) \ = \ f(z_0) + (z - z_0)f'(z_0) + \cdots .$$

We concentrate on the number $f'(z_0)$. This is called the *multiplier* of f at z_0, denoted by $\lambda_f(z_0)$. Multipliers are used to classify the fixed points into following categories:

1. z_0 is called an *attracting fixed point* if $0 < |\lambda_f(z_0)| < 1$;
2. z_0 is called a *repelling fixed point* if $|\lambda_f(z_0)| > 1$;
3. z_0 is called an *indifferent fixed point* or *neutral fixed point* if $|\lambda_f(z_0)| = 1$.

Further, a neutral fixed point z_0 is classified as

- a *rational* neutral fixed point if $\lambda^n = 1$ for some positive integer n and
- an *irrational* neutral fixed point if $\lambda^n \neq 1$ for any choice of positive integer n.

We now define the basin of attraction for a fixed point, say z_0.

Definition 6.1 The *basin of attraction* of a fixed point z_0 for a rational map R is defined to be the set of all points z such that $R^n(z) \to z_0$, as $n \to \infty$.

By virtue of the definition, it is clear that for any arbitrarily small $\epsilon > 0$, the basin of attraction of z_0 coincides with the union of the backward iterates $R^{-n}(D(z_0, \epsilon))$, where $D(z_0, \epsilon)$ is an ϵ-neighbourhood around the point z_0. Since the analysis is local, we consider the restriction $f = R : D(0, \epsilon) \longrightarrow \mathbb{C}$ such that $f(0) = 0$. Now, the Taylor series of R in a neighbourhood of 0 will be of the form

$$R(z) \ = \ R(0) + zR'(0) + \cdots \ = \ zR'(0) + \cdots \tag{8}$$

Case 1: Suppose 0 is an Attracting Fixed Point

Theorem 6.2 *There exists a unique injective holomorphic map*

$$\phi : D(0,\epsilon) \longrightarrow D(0,\epsilon)$$

such that the following diagram commutes:

$$\begin{array}{ccc} D(0,\epsilon) & \xrightarrow{R} & D(0,\epsilon) \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi} \\ D(0,\epsilon) & \xrightarrow{\lambda z} & D(0,\epsilon) \end{array}$$

i.e., $\phi^{-1} \circ R \circ \phi(z) = \lambda z$ *for* $\lambda = R'(0)$.

First proof. Consider the following diagram

$$\begin{array}{ccccccc} D(0,\epsilon) & \xrightarrow{R} & D(0,\epsilon) & \xrightarrow{R} & D(0,\epsilon) & \dashrightarrow^{R} & \cdots \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi} & & \\ D(0,\epsilon) & \xrightarrow{\lambda z} & D(0,\epsilon) & \xrightarrow{\lambda z} & D(0,\epsilon) & \dashrightarrow^{\lambda z} & \cdots \end{array}$$

Define

$$\phi(z) \;:=\; \lim_{n\to\infty} \frac{R^n(z)}{\lambda^n} \;=\; \lim_{n\to\infty} \phi_n(z), \quad \text{where } \phi_n(z) \;:=\; \frac{R^n(z)}{\lambda^n}.$$

Then,

$$\phi_{n+1}(z) \;=\; \frac{1}{\lambda^{n+1}} R^n \circ R(z) \;=\; \frac{1}{\lambda}\phi_n \circ R(z).$$

As $n \to \infty$, we have

$$\phi(z) \;=\; \frac{1}{\lambda}\phi(R(z)), \;\Longrightarrow\; \phi \circ R \circ \phi^{-1}(z) \;=\; \lambda(z),$$

meaning ϕ is a conjugation.

Note that for some small $\delta > 0$, we have $|R(z) - \lambda z| \le c|z|^2$ whenever $|z| < \delta$. Thus, $|R(z)| \le |\lambda||z| + c|z|^2 \le (|\lambda| + c\delta)|z|$. Further, by induction, we also have that

$$|R^n(z) \le (|\lambda| + c\delta)^n |z|, \text{ whenever we choose } \delta > 0 \text{ to satisfy } |\lambda| + c\delta < 1 \text{ and } |z| < \delta.$$

Consider

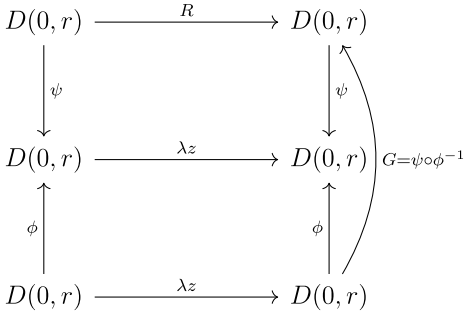
$$|\phi_{n+1}(z) - \phi_n(z)| = \left| \frac{R \circ R^n(z) - \lambda R^n(z)}{\lambda^{n+1}} \right| \leq \frac{c|R^n(z)|^2}{\lambda^{n+1}} \leq \frac{(|\lambda| + c\delta)^n |z|}{|\lambda|},$$

for $|z| < \delta$. Hence for sufficiently small δ, ϕ_n converges uniformly for all $|z| < \delta$, and the conjugation exists.

What now remains is to check the uniqueness of ϕ. Since $\phi_n(z) = \frac{1}{\lambda^n} R^n(z)$ and $\phi(0) = 0, \forall n \in \mathbb{Z}_+$, we make use of chain rule to find that $\phi'(0) = 1$. Thus ϕ is a local conformal map that is injective near 0. For some $r > 0$ that determines this neighbourhood, consider the following commutative diagram,

$$\begin{array}{ccc} D(0,r) & \xrightarrow{R} & D(0,r) \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi} \\ D(0,r) & \xrightarrow{z \mapsto \lambda z} & D(0,r) \end{array}$$

Suppose $\psi : D(0,r) \longrightarrow D(0,r)$ is another injective holomorphic map such that the following diagram commutes.



Then $\lambda G(z) = G(\lambda z)$, where $G = \psi \circ \phi^{-1}$. This implies

$$\lambda(c_1 z + c_2 z^2 + \cdots) = c_1(\lambda z) + c_2(\lambda z)^2 + \cdots \quad (\text{since } G(0) = c_0 = 0).$$

Then, by comparing the coefficients, we get,

$$c_2\lambda = c_2\lambda^2; \;\; \cdots \;\; c_n\lambda = c_n\lambda^n; \;\; \cdots .$$

However, since $0 < |\lambda| < 1$, we get $c_2 = c_3 = \cdots = 0$. Also $G'(0) = 1$ implies $c_1 = 1$. Hence, $\psi \circ \phi^{-1}(z) = G(z) = z$, thereby making ϕ to be unique. □

Second proof. Here, we only prove the formal existence of a power series. We appeal to the first proof for the convergence of the power series.

We need to justify the equation $\phi^{-1} \circ R \circ \phi(z) = \lambda z$ where $\lambda = R'(0)$, $0 < |\lambda| < 1$. This is equivalent to studying the functional equation $R \circ \phi(z) = \phi(\lambda z)$ for some ϕ.

We assume that ϕ has a formal power series expansion, $\phi(z) = \sum_{j=0}^{\infty} c_j z^j$, where $\phi(0) = 0$ i.e., $c_0 = 0$. Also we have $R(z) = \lambda z + O(z^2)$. So,

$$R(\phi(z)) = \lambda\phi(z) + O(\phi(z)^2) = \phi(\lambda z) = \sum_{j=0}^{\infty} c_j(\lambda z)^j.$$

In other words,

$$\lambda(c_1 z + c_2 z^2 + \cdots) + \sum_{j=2}^{\infty} \alpha_j(\phi(z))^j = \lambda z + \lambda^2 c_2 z^2 + \cdots.$$

Substituting for $\phi(z)$ and by comparing the coefficients, we obtain,

$$\begin{aligned} \lambda c_1 &= \lambda c_1; \\ \lambda c_2 + \alpha_2 c_1^2 &= \lambda^2 c_2; \\ \lambda c_3 + \alpha_2(c_1 z + c_2 z^2 + \cdots)^2 \alpha_2(c_1 z + \cdots)^3 &= \lambda^3 c_3; \\ &\cdots \end{aligned}$$

that implies

$$\begin{aligned} c_2 &= \frac{\alpha_2 {c_1}^2}{\lambda^2 - \lambda}; \\ c_3 &= \frac{2\alpha_2 c_1 c_2 + \alpha_3 c_1^3}{\lambda^3 - \lambda}; \\ &\cdots. \end{aligned}$$

Since $|\lambda| < 1$, all the coefficients are well defined. Thus, given c_1, we can inductively evaluate and find the coefficients. Hence, the functional equation $R \circ \phi(z) = \phi(\lambda z)$ has a well defined solution. □

Case 2: Suppose 0 is a repelling fixed point. As in Case 1, a similar local analysis holds viz. there exists a unique injective holomorphic map $\phi : D(0, r) \longrightarrow D(0, r)$ which makes the following diagram commutative,

$$\begin{array}{ccc} D(0,r) & \xrightarrow{R} & D(0,r) \\ \uparrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \phi} \\ D(0,r) & \xrightarrow{z \mapsto \lambda z} & D(0,r) \end{array}$$

i.e., $\phi^{-1} \circ R \circ \phi(z) = \lambda z \implies \phi^{-1} \circ R^n \circ \phi(z) = \lambda^n z$ as $|\lambda| > 1$. Hence, $|\lambda|^n \to \infty$ as $n \to \infty$. Thus, $\{R^n : n \geq 1\}$ does not form an equicontinuous family and so the repelling fixed points are in J_R whereas the attracting fixed points are in F_R.

Case 3: Suppose 0 is an indifferent fixed point. Here, we have two sub-cases depending on whether 0 is a rationally indifferent fixed point i.e, whether λ is a root of unity or if 0 is an irrationally indifferent fixed point i.e, if $\lambda = e^{2\pi i\theta}$, where θ is an irrational no.

Case 3a: Suppose 0 is a rationally indifferent fixed point. In this case, write

$$R(z) = az + bz^r + \cdots$$

near $z = 0$, where $a^m = 1$ for some positive integer m and $b \neq 0$. Then

$$R^m(z) = z + cz^r + \cdots$$

near the origin and $c \neq 0$ (as otherwise all the higher order derivatives of R at the origin will vanish forcing R^m to be the identity which is not possible as its degree is at least 2). The iterates of R^m will all be of the form

$$R^{mn}(z) = z + ncz^r + \cdots$$

and this means that the rth derivative of R^{mn} at the origin diverge to infinity and this forces $0 \in J_R$. A complete local analysis of the iterates of R near the origin requires many steps and upshot of it all is that R is not locally conjugate to its derivative at the origin near it. To provide a flavour of what is involved, let us look at

$$P(z) = z - z^2$$

that has a rationally indifferent fixed point at the origin. If z approaches the origin along the x-axis, then $0 < P(x) < x$ for $x \in (0, 1)$ and hence $P^n(x) \to 0$ as $n \to \infty$. In fact, the convergence to the origin holds on each small disc centered at $r \in (0, 1)$ of radius r. The union of these discs is a bulb (or a petal) containing a segment on the positive x-axis and the origin on its boundary. In general, if

$$R(z) = z + az^{p+1} + \cdots$$

near the origin, there exist p distinct petals each containing the origin on its boundary such that the iterates of R on each petal converge to the fixed point (i.e., the origin) in such a way that the argument of R^n approaches a fixed multiple of $2\pi/p$. The details of the proof can be found in [2, 7] for example.

Case 3b: Suppose 0 is an irrationally indifferent fixed point. Suppose $\lambda = e^{2\pi i\theta}$, where θ is an irrational number.

Definition 6.3 A real number θ is *Diophantine* if it is badly approximable by rational numbers, i.e., there exists a $c > 0$ so that

$$\left|\theta - \frac{p}{q}\right| \geq \frac{c}{q^k} \ \forall p, q \in \mathbb{Q} \text{ and } \forall k \geq 2.$$

Siegel proved that if $\lambda = e^{2\pi i\alpha}$, where α is a Diophantine number, then, there exists a conjugacy ϕ such that $\phi^{-1} \circ R \circ \phi(z) = \lambda z$ i.e., R is an irrational rotation of unit disc $\mathbb{D}$. The Fatou component containing the irrationally indifferent fixed point is called a *Siegel disc*.

7 Brolin's Theorem

It can shown that (see Theorem (4.2.7) in [2] for example) that if P is a polynomial with degree at least 2 and $z \in J_P$ is an arbitrary point, then the backward orbit

$$O^-(z) = \{w \ : \ \text{there exists } n \geq 0 \text{ such that } P^n(w) = z\}$$

is dense in J_P. Brolin's theorem quantifies this statement in a rather precise way using the notion of an equilibrium measure of a compact set. To briefly describe what this is, we start with the definition of a subharmonic function.

Definition 7.1 A function u on a domain $\Omega \subset \mathbb{C}$ taking values in $[-\infty, \infty)$ is said to be *subharmonic* if

(i) u is upper semi-continuous on Ω, i.e., $\limsup_{z\to p} u(z) \leq u(p)$ for all $p \in \Omega$, and
(ii) for each $p \in \Omega$, there exists $r_0 > 0$ such that

$$u(p) \leq \frac{1}{2\pi} \int_0^{2\pi} u(p + re^{i\theta})\, d\theta$$

for all $r \leq r_0$.

Harmonic functions provide smooth examples of subharmonic functions, but more illustrative examples are provided by looking at $u(z) = \log|f(z)|$ for a holomorphic function f on Ω. It can be shown that a C^2-smooth function u is subharmonic if and only if $\Delta u \geq 0$ on Ω.

Definition 7.2 A set E is said to be *polar set* if there is a subharmonic function u such that $E \subset \{u = -\infty\}$.

For example, the zero set of a holomorphic function $f(z)$ is a polar set. Since, then $\log(|f(z)|)$ is subharmonic and $f(z) = 0$ precisely when $\log|f(z)| = -\infty$. Next, we define potential and energy.

Definition 7.3 Let μ be a finite Borel measure on $\mathbb{C}$ such that support of μ is compact. Then the *potential* $p_\mu : \mathbb{C} \longrightarrow [-\infty, \infty]$ is defined as

$$p_\mu(z) \ = \ \int_{\mathbb{C}} \log|z - w|\, d\mu(w).$$

It can be shown that p_μ is subharmonic on $\mathbb{C}$ and harmonic away from the support of μ.

Definition 7.4 The *energy* associated to the potential p_μ is defined as

$$I_\mu \;=\; \int_{\mathbb{C}} p_\mu(z)\, d\mu(z).$$

To define the equilibrium measure of a compact set $K \subset \mathbb{C}$, let $\mathcal{P}(K)$ be the collection of all Borel probability measures on K that are supported on K. If there exists $\nu \in \mathcal{P}(K)$, such that

$$\sup_{\mu \in \mathcal{P}(K)} I_\mu \;=\; I_\nu,$$

then ν is called the *equilibrium measure* for K. Note that ν always exists because every sequence in $\mathcal{P}(K)$ admits a weakly convergent subsequence. Furthermore, ν is unique, if K is non-polar. A theorem of Frostman shows that

$$p_\nu(z) \le I_\nu,$$

everywhere on $\mathbb{C}$ and that equality holds on $K \backslash E$ where E is a F_σ polar subset of the boundary ∂K.

Theorem 7.5 (Brolin's theorem) *Let $P(z)$ be polynomial of degree d and J_P be its Julia set. Start with any $w \in J_P$. Then $P^n(z) = w$ has d^n roots with counting multiplicity. Define,*

$$\mu_n \;= \frac{1}{d^n} \sum_{p^n(\xi)=w} \delta_\xi, \qquad \textit{where} \qquad \delta_\xi(E) = \begin{cases} 1 & \textit{if } \xi \in E \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\mu_n \in \mathcal{P}(J_P)$ and μ_n converges weakly to ν. In fact, ν is the equilibrium measure of J_P.

A proof can be found in [8]. The measure ν also has the following property:

Theorem 7.6 *Let $P(z)$ be as above. Then entropy of the polynomial P with respect to the measure ν is* $\log d$. *ν is unique measure of maximal entropy.*

8 What Happens in Higher Dimensions?

Moving from the plane to $\mathbb{C}^2$, a natural class of mappings to focus on are polynomial automorphisms, i.e., $P = (P_1, P_2) : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ such that each component is a polynomial. It is a theorem that if such an P is both injective and surjective, then its inverse

is also a polynomial. As on the plane, the goal is to understand the behaviour of the iterates of P. A theorem of Milnor-Friedland lists all possible conjugacy classes in the group of polynomial automorphisms of $\mathbb{C}^2$ and identifies the class of generalized Hénon mappings as one which is dynamically interesting. A generalized Hénon map is defined as

$$H(x, y) = (y, P(y) - \delta x)$$

where P is monic polynomial in y of degree $d \geq 2$ and $\delta \neq 0$. We may consider finite compositions of such maps and understand their iterates. As a first step, the dynamics of H can be understood by observing that the degree of the second component of the iterates of H is always larger than the first. It follows that there is a filtration of $\mathbb{C}^2$ in each component of which the iterates of H have a well defined behaviour.

Thus, define $K^{\pm}$ to be the set with bounded forward/backward orbits, $K = K^+ \cap K^-$, $J^{\pm} = \partial K^{\pm}$, $J = J^+ \cap J^-$ and $U^+ = \mathbb{C}^2 \setminus K^+$. It can be shown that the family of iterates H^n is normal on the interior of K^+ (if it is non-empty) and that in no neighbourhood of any point in J^+ is this true. Furthermore, the Fatou set of H is exactly $\mathbb{C}^2 \setminus J^+$. The absence of a useful analog of Montel's theorem forces different tools to be brought into the picture at this stage and it turns out that methods from pluripotential theory have been very useful. The Green's function

$$G(x, y) = \lim_{n \to \infty} d^{-n} \log^+ |H^n(x, y)|$$

exists as a non-negative plurisubharmonic function on $\mathbb{C}^2$ that vanishes precisely on K^+ and which is pluriharmonic away from it. The definition also implies that $G^+ \circ H = dG^+$ everywhere. Similarly, we can define G^- by considering H^{-1} and its iterates. Since both $G^{\pm}$ are plurisubharmonic,

$$\mu^{\pm} = \frac{1}{2\pi} dd^c G^{\pm}$$

are well defined positive closed currents whose supports are precisely $J^{\pm}$. These currents have a laminar structure in the measure theoretic sense and many of their properties hinge upon this fact. There is a version of Brolin's theorem as well in which points are replaced by the pullbacks of a small disc. Convergence to a positive multiple of μ^+ is then the conclusion. The measure $\mu = \mu^+ \wedge \mu^-$ is well defined and invariant under H. It is a theorem that μ is mixing and hence ergodic. Furthermore, the entropy of H is $\log d$.

These ideas have been developed for holomorphic endomorphisms of $\mathbb{P}^n$ as well—see [4] for a comprehensive survey.

References

1. Ahlfors, L. V. (1953). *Complex analysis: An introduction to the theory of analytic functions of one complex variable*. New York-Toronto-London: McGraw-Hill Book Company Inc.
2. Beardon, A. F. (1991). *Iteration of rational functions: Complex analytic dynamical systems*, Graduate texts in mathematics (Vol. 132). New York: Springer.
3. Bedford, E., & Smillie, J. (1991). Polynomial diffeomorphisms of $\mathbb{C}^2$: Currents, equilibrium measure and hyperbolicity. *Inventiones Mathematicae, 103*(1), 69–99.
4. Dinh, T. -C., & Sibony, N. (2010). Dynamics in several complex variables: Endomorphisms of projective spaces and polynomial-like mappings (English summary), *Holomorphic dynamical systems* (pp. 165–294), Lecture notes in mathematics (1998). Berlin: Springer.
5. Fornaess, J. E., & Sibony, N. (1992). Complex Hénon mappings in $\mathbb{C}^2$ and Fatou-Bieberbach domains. *Duke Mathematical Journal, 65*(2), 345–380.
6. Hubbard, J. H., & Oberste-Vorth, R. W. (1994). Hénon mappings in the complex domain: I. The global topology of dynamical space. *Institute des Hautes Études Scientifiques Publications Mathématiques, 79,* 5–46.
7. Milnor, J. (2006). *Dynamics in one complex variable* (3rd ed.). Annals of mathematics studies (Vol. 160). Princeton: Princeton University Press.
8. Ransford, T. (1995). *Potential theory in the complex plane*, London mathematical society student texts (Vol. 28). Cambridge: Cambridge University Press.
9. Steinmetz, N. (1993). *Rational iteration: Complex analytic dynamical systems*, de Gruyter studies in mathematics (Vol. 16). Berlin: Walter de Gruyter & Co.

Topics in Homogeneous Dynamics and Number Theory

Anish Ghosh

1 Introduction

This is a survey of some topics at the interface of dynamical systems and number theory, based on lectures delivered at CIRM Luminy, the University of Houston, and IIT Delhi. Specifically, we will be interested in the ergodic theory of group actions on homogeneous spaces and its connections to metric Diophantine approximation. The topics covered in the lectures included the study of the Diophantine approximation of linear forms using dynamics, the study of quadratic forms in particular the famous Oppenheim's conjecture and its variations, as well as lattice point counting using dynamics. At IIT, non-divergence estimates for unipotent flows and Margulis' proof of the Borel Harish-Chandra theorem using the non-divergence estimates were also covered. There are many recent and excellent surveys covering all this material, including but not restricted to [9, 24, 27, 49, 50, 60] and the books [23, 59]. Rather than reinvent the wheel, I have chosen to present some other recent topics at the interface of Diophantine approximation and homogeneous dynamics in this article. The topics chosen are representative of the lectures but reflect my interests and problems that I have been recently involved with. While some of the lectures were at a more basic level, this article serves as an introduction to more advanced and more recent material. In particular, this is not meant to be a comprehensive survey of this very active and rapidly expanding subject, a shortcoming redressed by the many aforementioned surveys. The hope is that this article will serve as a guide for students with some preparation, e.g. the ones who attended the lectures and point them to further reading and interesting research avenues. Two sections of this article are devoted to results using methods from classical number theory, an indispensable part of the toolbox of anyone interested in Diophantine analysis.

A. Ghosh (✉)
School of Mathematics, Tata Institute for Fundamental Research, Mumbai, India
e-mail: ghosh@math.tifr.res.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_6

1.1 Homogeneous Dynamics

Let G be a unimodular, locally compact, second countable topological group and Γ be a lattice in G. The homogeneous space G/Γ is equipped with a finite measure which descends from the Haar measure on G and which can therefore be normalised to make G/Γ a probability space. A subgroup H of G acts on the probability space G/Γ by translations. The ergodic theory of this action has been extensively studied in recent decades, and is referred to as homogeneous dynamics. For particular choices of G, H and Γ, the spaces G/Γ and $H\backslash G$ parametrise objects of number theoretic interest in many cases and the ergodic theory of the H action on G/Γ (resp. the Γ action on $H\backslash G$) gives valuable Diophantine information about these objects. Here are some examples:

Example 1.1 $G = \mathrm{SL}_n(\mathbb{R})$ and $\Gamma = \mathrm{SL}_n(\mathbb{Z})$. Then G/Γ can be identified with the space of unimodular lattices in $\mathbb{R}^n$. The dynamics of diagonal flows on G/Γ plays an important role in Diophantine approximation of vectors and linear forms, as explained in the next section.

Example 1.2 Let $n = p + q$ and set $H = \mathrm{SO}(p, q)$. Then the H action on $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$ plays an important role in the study of quadratic forms. This is also touched upon in the next section.

Example 1.3 $G = \mathrm{SL}_n(\mathbb{R}) \times \mathrm{SL}_n(\mathbb{Q}_p)$ and $\Gamma = \mathrm{SL}_n(\mathbb{Z}[1/p])$. Once again, Γ is a non-cocompact lattice in G and similar to the example above, one can identify G/Γ can be identified with the space of discrete $\mathbb{Z}[1/p]$-modules in $\mathbb{R}^n \times \mathbb{Q}_p^n$. Dynamics on this and related spaces plays an important role in p-adic Diophantine approximation.

Example 1.4 Let k be a degree d number field, S be the set of Archimedean places and O_k be it's ring of integers. Then O_k is a lattice in $k_S := \mathbb{R} \times \cdots \times \mathbb{R} \times \mathbb{C} \times \cdots \times \mathbb{C}$ via the Galois embedding. By the Borel Harish-Chandra theorem, $\Gamma = \mathrm{SL}_2(O_k)$ is a lattice in $\mathrm{SL}_n(k_S) := \prod_{s\in S} \mathrm{SL}_n(k_s)$ where k_s denotes the completion of k at the place $s \in S$. In Sect. 3, we will consider flows on the space G/Γ and connections to Diophantine approximation in number fields. Such dynamics is intimately connected to geodesic flows on the associated arithmetic orbifold.

1.2 Diophantine Approximation

Diophantine approximation begins with a theorem due to Dirichlet [16] which states that

Theorem 1.5 *For any $x \in \mathbb{R}$ and any $Q > 0$, there exist $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that*

$$|qx - p| \quad < \quad \frac{1}{Q} \text{ and } q \leq Q.$$

As a corollary, it follows that for every $x \in \mathbb{R}$, there exist infinitely many $q \in \mathbb{N}$ such that

$$|qx - p| \quad < \quad \frac{1}{q},$$

for some $p \in \mathbb{Z}$. One can also consider Diophantine approximation in higher dimension. Indeed, the role of homogeneous dynamics is brought into sharper focus in higher dimensions as it serves as a replacement for continued fractions, an efficient theory of which is only available in dimension 1. For instance, the corollary to Dirichlet's theorem in arbitrary dimension reads as follows:

Theorem 1.6 *For every* $\mathbf{x} \in \mathbb{R}^d$*, there exist infinitely many* $q \in \mathbb{Z}$ *such that*

$$\|q\mathbf{x} - \mathbf{p}\| < |q|^{-1/d}, \tag{1}$$

for some $\mathbf{p} \in \mathbb{Z}^d$.

Here, $\|\cdot\|$ is the supremum norm. In dimension $d > 1$, there are two possible settings for Diophantine approximation. The above is called the simultaneous setting, one could also consider the *dual* setting where one considers small values of the linear form

$$|\mathbf{q} \cdot \mathbf{x} + p|$$

for $\mathbf{q} \in \mathbb{Z}^d$ and $p \in \mathbb{Z}$. The simultaneous and dual settings are related by *transference principles*. We refer the reader to [16] for Khintchine's classical transference principles and [19] for some recent developments involving transference inequalities in the weighted and inhomogeneous settings. Dirichlet's theorem can be proved using the pigeonhole principle (as proved originally by Dirichlet) and also using Minkowski's theorem in the geometry of numbers.

The next major theorem in metric Diophantine approximation seeks to expand the class of approximating functions. Let ψ be a non-increasing function from $\mathbb{R} \longrightarrow \mathbb{R}_+ \cup \{0\}$ be given and let $\mathcal{W}_d(\psi, \mathbb{R})$ be the subset of real numbers x for which there exist infinitely many $\mathbf{q} \in \mathbb{Z}^d$ such that

$$|\mathbf{q} \cdot \mathbf{x} + p| < \psi(\|\mathbf{q}\|^d) \tag{2}$$

for some $p \in \mathbb{Z}$. Khintchine's theorem (in dual form) characterizes the size of $\mathcal{W}_d(\psi, \mathbb{R})$ in terms of Lebesgue measure.

Theorem 1.7 (Khintchine's theorem) $\mathcal{W}_d(\psi, \mathbb{R})$ *has zero or full measure according as*

$$\sum_{x=1}^{\infty} \psi(x) \tag{3}$$

converges or diverges.

There exist numbers (and vectors) for which (1) cannot be improved; these are called badly approximable. We now define the same.

Definition 1.8 A vector $\mathbf{x} \in \mathbb{R}^d$ is called *badly approximable* if there exists $c := c(\mathbf{x}) > 0$ such that

$$\|q\mathbf{x} - \mathbf{p}\| \geq \frac{c}{|q|^{1/d}}. \tag{4}$$

It is well known that badly approximable vectors have zero Lebesgue measure and full Hausdorff dimension (Jarnik [45] for $n = 1$ and Schmidt [65, 66] for arbitrary n). In fact, Schmidt showed that they are *winning* for a certain game, a stronger and more versatile property than having full Hausdorff dimension. We refer the reader to Dani's article [21] in this volume for an introduction to Schmidt's game.

On the opposite end of the spectrum to badly approximable vectors, are singular vectors.

Definition 1.9 A vector $\mathbf{x} \in \mathbb{R}^d$ is said to be *singular* if for every $\varepsilon > 0$ there exists N_0 with the following property: for each $N \geq N_0$, there exist $\mathbf{p} \in \mathbb{Z}^d$, $q \in \mathbb{N}$ so that

$$\|q\mathbf{x} - \mathbf{p}\| < \frac{\varepsilon}{N^{\frac{1}{d}}} \quad \text{and} \quad q < N. \tag{5}$$

In other words, $\mathbf{x}$ is singular if Dirichlet's Theorem can be improved by an arbitrarily small constant factor $\varepsilon > 0$. In the case $d = 1$, Khintchine [48] showed that a real number is singular if and only if it is rational. Moreover, it was shown by Davenport & Schmidt [22] that the set of singular vectors has zero Lebesgue measure.

We now move from linear forms to quadratic forms and briefly discuss Oppenheim's conjecture. Let $n \geq 3$ and let Q be a non degenerate indefinite quadratic form in n variables and assume that Q is not proportional to a form with rational coefficients. It was a conjecture of Oppenheim from the 1920s and a celebrated theorem of Margulis [56] that under these conditions $Q(\mathbb{Z}^n)$ is dense in $\mathbb{R}$. Oppenheim's conjecture is false for binary quadratic forms, a counterexample can be constructed using badly approximable numbers. More precisely, the quadratic form $Q(x, y) = y^2 - \theta^2 x^2$ where θ is a quadratic irrational with θ^2 irrational provides a counterexample. For this and more, we refer the reader to Borel's survey [12].

How does the Diophantine approximation of vectors and linear and quadratic forms relate to dynamics of subgroup actions on G/Γ or lattice actions on $H\backslash G$? Let $G = \mathrm{SL}_{n+1}(\mathbb{R})$ and $\Gamma = \mathrm{SL}_{n+1}(\mathbb{Z})$, then G/Γ and can be naturally identified with the space Ω_{n+1} of unimodular, i.e., covolume 1 lattices in $\mathbb{R}^{n+1}$. Namely, G acts transitively on Ω_{n+1} by multiplication and the stabilizer of the lattice $\mathbb{Z}^{n+1}$ is $\mathrm{SL}_{n+1}(\mathbb{Z})$. The space G/Γ is a non-compact, finite volume space and Mahler's compactness criterion describes the compact subsets of G/Γ. Diophantine approximation of vectors in $\mathbb{R}^n$ can be modelled using dynamics of subgroup actions on Ω_{n+1}. Given a vector $\mathbf{x} \in \mathbb{R}^n$, we consider the unimodular lattice

$$\Lambda_{\mathbf{x}} := \begin{pmatrix} 1 & \mathbf{x} \\ 0 & \mathrm{Id} \end{pmatrix} \mathbb{Z}^{n+1} \in \Omega_{n+1}.$$

Further, let

$$g_t := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \in G.$$

Then we have the following two propositions connecting Diophantine properties of $\mathbf{x}$ with the dynamics of the g_t action on G/Γ due to Dani [20]. The first concerns badly approximable vectors.

Proposition 1.10 *A vector* $\mathbf{x} \in \mathbb{R}^n$ *is badly approximable if and only if* $\{g_t \Lambda_{\mathbf{x}} : t > 0\}$ *is bounded in* G/Γ.

And the second concerns singular vectors.

Proposition 1.11 *A vector* $\mathbf{x} \in \mathbb{R}^n$ *is singular if and only if* $\{g_t \Lambda_{\mathbf{x}} : t > 0\}$ *is divergent in* G/Γ.

Kleinbock and Margulis [52] have proved a more general version of the "Dani correspondence" and have provided a dynamical proof of Khintchine's theorem using exponential mixing of the g_t action on G/Γ. This is closely related to the *shrinking target problem* for group actions on homogeneous spaces. In [68], Sullivan established the following folklore theorem. Let $V = \mathbb{H}^{d+1}/\Gamma$ be a hyperbolic manifold where Γ is a discrete subgroup of hyperbolic isometries which is not co-compact, and let $\operatorname{dist} v(t)$ denote the distance from a fixed point in V of the point achieved after traveling a time t along the geodesic with initial direction v.

Theorem 1.12 ([68]) *For all* $x \in V$, *and almost every* $v \in T_x V$,

$$\limsup_{t \to \infty} \frac{\operatorname{dist} v(t)}{\log t} = \frac{1}{d}.$$

Kleinbock and Margulis generalized Sullivan's logarithm law to locally symmetric spaces. In fact, both the logarithm law and Khintchine's theorem are manifestations of a 0–1 Borel–Cantelli type law for diagonal flows on homogeneous spaces. This scheme was subsequently carried out in the positive characteristic setting in [4] (see also [34]), and in the setting of geodesic orbits on the frame bundle of finite volume non-compact hyperbolic manifolds in [7]. We refer the reader to [33] for an introduction to the shrinking target problem, including a more general formulation of the logarithm law for (not necessarily one-parameter") subgroups, as well as a list of references.

How does Oppenheim's conjecture relate to dynamics? Let $G = \mathrm{SL}_n(\mathbb{R})$, $H = \mathrm{SO}(p, q)$ and $\Gamma = \mathrm{SL}_n(\mathbb{Z})$. The space $H\backslash G$ can be identified with the space of quadratic forms of signature (p, q) in $n = p + q$ variables. The relevant dynamics here is that of the Γ action on $H\backslash G$ (or, dually, the H action on G/Γ). More precisely, the following proposition implies Oppenheim's conjecture.

Proposition 1.13 *Any H orbit on G/Γ is either closed and carries an H° invariant measure or is dense.*

The above result was proved by G. Margulis [56] for ternary quadratic forms, thus settling Oppenheim's conjecture. The main point is that, under the conditions above, $\mathrm{SO}(p, q)$ is generated by unipotent one-parameter subgroups. The proposition above is an instance of the conjectures of Raghunathan and Dani on orbit closures and invariant measures for actions of such groups on homogeneous spaces. These conjectures were settled by M. Ratner. We refer the reader to [59] for details on this beautiful subject. Recently, the Γ action on $H\backslash G$ has been used to study *effective* versions on Oppenheim's conjecture. We will not elaborate on this theme in this survey, referring the reader instead to [39] for details and to [37, 38] for the more general study of effective density of lattice orbits on homogeneous varieties and [35, 36] for the related problem of intrinsic Diophantine approximation on varieties. See also [32] for a related question on quadratic forms studied originally by Bourgain [13].

Structure. The rest of the article is divided into three sections. The next section focuses on Dirichlet's theorem and considers two different aspects—probabilistic and geometric, of the problem of *distribution of approximates*. The tools in this section involve equidistribution of flows and limiting distributions for flows on the space of lattices. Section 3 considers Diophantine approximation in number fields. The classical theorems like Dirichlet's theorem and Khintchine's theorem can be generalised to number fields. We pay particular attention to vectors which are badly approximable by rationals from a number field and the associated dynamics of diagonal flows on arithmetic orbifolds. The last two sections discuss Diophantine approximation in two diverse settings: that of projective space and hyperbolic space. In Sect. 4, we present a projective version of Khintchine's theorem and the more general Duffin–Schaeffer conjecture and in Sect. 5, we discuss Diophantine approximation by orbits of Fuchsian and Kleinian groups on the boundary of hyperbolic space. The techniques used in the last two sections are classical in nature.

2 On the Distribution of Approximates

In this section, we describe some recent results on the distribution of approximates in Dirichlet's theorem. First, we describe a probabilistic distribution problem, originally due to Erdős, Szüsz and Turán and developed in a homogeneous context, in [3]. In the next subsection, we consider the geometry of the approximates and describe a spiraling equidistribution of approximates proved in [6].

2.1 The EST Distribution

In 1958, Erdős, Szüsz and Turán [26] introduced a problem in probabilistic Diophantine approximation: what is the probability $f(N, A, c)$ that a point α chosen from the uniform distribution on [0, 1] has a solution $\frac{p}{q} \in \mathbb{Q}$ to the inequality

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{A}{q^2} \tag{6}$$

with the constraint that the denominator q lies in $[N, cN]$? Here $A > 0$, $c > 1$ are fixed positive parameters, and N is a parameter which goes to infinity. The above inequality is of course a close variant of the inequality in Dirichlet's theorem. In particular, we know that $A = 1$ admits infinitely many solutions and that (by Hurwitz's theorem), $A = \frac{1}{\sqrt{5}}$ is the best allowable constant which admits infinitely many solutions for all α. Given A, c, N, let $\mathrm{EST}(A, c, N)(\alpha)$ be the number of solutions $p/q \in \mathbb{Q}$ with $\gcd(p, q) = 1$ to (6). Letting $\alpha \in [0, 1]$ be a uniform random variable yields an integer-valued random variable $\mathrm{EST}(A, c, N)$, with

$$P\left(\mathrm{EST}(A, c, N) = k\right) \;=\; m\left(\alpha \in [0, 1] \;:\; \text{there are exactly } k \text{ solutions to (6)}\right),$$

where m is Lebesgue measure on [0, 1]. Then, the EST question is the existence of the limit

$$\lim_{N \to \infty} P\left(\mathrm{EST}(A, c, N) > 0\right).$$

Kesten [46] considered a modified version of this problem, he defined the sequence of random variables $K(A, N)$ as the number of solutions to

$$|\alpha q - p| \;\leq\; \frac{A}{N}, \quad 1 \leq q \leq N, \tag{7}$$

where α is a uniform [0, 1] random variable. That is,

$$P(K(A, N) = k) = m\left(\alpha \in [0, 1] \;:\; \text{there are exactly } k \text{ solutions to (7)}\right).$$

The Kesten distribution was studied by Marklof [Theorem 4.4 in [57]], thought at the time he was not aware of Kesten's question. In [3], it was shown that the limiting distributions of the random variables $\mathrm{EST}(A, c, N)$ and $K(A, N)$ exist as $N \to \infty$. In fact, they can be viewed as the probability of a random unimodular lattice intersecting a certain fixed region. Let μ_2 denote the Haar probability measure on Ω_2, the space of unimodular lattices in $\mathbb{R}^2$. Given $\Lambda \in X_2$, $\Lambda = g\mathbb{Z}^2$, let Λ_{prim} be the set of primitive vectors in Λ. It should be emphasized that we were unaware of Marklof's work, and the Kesten part of the theorem below merely reproves a part of Marklof's theorem referred to above.

Theorem 2.1 ([3] Theorem (1.1)) *The limiting distribution of the random variables* $\mathrm{EST}(A, c, N)$ *and* $K(A, N)$ *exist and denoting the random variables with these limiting distributions as* $\mathrm{EST}(A, c)$ *and* $K(A)$*, we have*

$$P(\mathrm{EST}(A, c) = k) = \mu_2(\Lambda \in X_2 \ : \ \#(\Lambda_{\mathrm{prim}} \cap H_{A,c}) = k), \tag{8}$$

and

$$P(K(A) = k) = \mu_2(\Lambda \in X_2 \ : \ \#(\Lambda_{\mathrm{prim}} \cap R_A) = k) \tag{9}$$

where

$$H_{A,c} = \{(x, y) \in \mathbb{R}^2 \ : \ xy \leq A, \ 1 \leq y \leq c\}, \tag{10}$$

and

$$R_A = \{(x, y) \in \mathbb{R}^2 \ : \ |x| \leq A, \ 0 \leq y \leq 1\}. \tag{11}$$

In fact, the setting in [3] is abstract and axiomatic, allowing for a great deal of flexibility. The philosophy of *equivariant point processes*, introduced in this context in [3], allows us to obtain the existence of these limiting distributions in higher dimensions, for linear forms, for points on smooth curves as well as in the setting of the set of holonomy vectors of saddle connections on translation surfaces. For details as well as a more detailed history of the problem, the reader is referred to [3]. Further applications of equivariant point processes are explored in a forthcoming monograph of Athreya and Ghosh.

2.2 Spiraling of Approximates

In this section, we describe some results from [6] (see also [5]) where the geometric study of the distribution of approximates in Dirichlet's theorem was initiated. We consider a vector $\mathbf{x} \in \mathbb{R}^d$ and form, as before, form the associated unimodular lattice in $\mathbb{R}^{d+1}$.

$$\Lambda_{\mathbf{x}} := \begin{pmatrix} \mathrm{Id}_d & \mathbf{x} \\ 0 & 1 \end{pmatrix} \mathbb{Z}^{d+1} = \left\{ \begin{pmatrix} q\mathbf{x} - \mathbf{p} \\ q \end{pmatrix} \ : \ \mathbf{p} \in \mathbb{Z}^d, q \in \mathbb{Z} \right\}.$$

Then we can view the approximates $(\mathbf{p}, q)$ of $\mathbf{x}$ appearing in (1) as points of the lattice $\Lambda_{\mathbf{x}}$ in the region

$$R := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^d \times \mathbb{R} \ : \ \|\mathbf{v}_1\| |v_2|^{1/d} \leq 1 \right\}. \tag{12}$$

The goal is to understand the geometry of the set of approximates $\Lambda_{\mathbf{x}} \cap R$. To do so, consider the following sets:

$$R_{\epsilon,T} := \{\mathbf{v} \in R \ : \ \epsilon T \leq v_2 \leq T\} \tag{13}$$

and, for a measurable subset A of $\mathbb{S}^{d-1}$ with zero measure boundary,

$$R_{A,\epsilon,T} := \left\{ \mathbf{v} \in R_{\epsilon,T} \ : \ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \in A \right\}. \tag{14}$$

For a unimodular lattice Λ, define

$$N(\Lambda, \epsilon, T) \ = \ \#\{\Lambda \cap R_{\epsilon,T}\}$$

and

$$N(\Lambda, A, \epsilon, T) \ = \ \#\{\Lambda \cap R_{A,\epsilon,T}\}.$$

Let dk denote Haar measure on $K = \mathrm{SO}_{d+1}(\mathbb{R})$, and let $X_{d+1} = \mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$ denote the space of unimodular lattices in $\mathbb{R}^{d+1}$. The following equidistribution theorem is proved in [6].

Theorem 2.2 *For every $\Lambda \in X_{d+1}$, $A \subset \mathbb{S}^{d-1}$ as above, and for every $\epsilon > 0$,*

$$\lim_{T\to\infty} \frac{\int_K N(k^{-1}\Lambda, A, \epsilon, T)\, \mathrm{d}k}{\int_K N(k^{-1}\Lambda, \epsilon, T)\, \mathrm{d}k} = \mathrm{vol}(A). \tag{15}$$

In other words, for any lattice Λ, on average over the set of directions $\mathbf{v}$, the set of approximates satisfying Dirichlet's theorem in the direction $\mathbf{v}$ equidistributes in the set of directions $\mathbb{S}^{d-1}$ in the orthogonal complement to $\mathbf{v}$. The proof of the Theorem 2.2 depends on an equidistribution result for Siegel transforms which is likely to have other applications.

One can also fix the vertical, and instead average the counting functions over a range of heights, T, to obtain a result for almost every lattice Λ, with respect to the probability measure μ on X_{d+1} induced by Haar measure on $\mathrm{SL}_{d+1}(\mathbb{R})$.

Theorem 2.3 *Fix $A \subset \mathbb{S}^{d-1}$ as above. For μ-almost every $\Lambda \in X_{d+1}$ and for every $\epsilon > 0$,*

$$\lim_{S\to\infty} \frac{\int_0^S N(\Lambda, A, \epsilon, e^t)\, \mathrm{d}t}{\int_0^S N(\Lambda, \epsilon, e^t)\, \mathrm{d}t} = \mathrm{vol}(A). \tag{16}$$

That is, if we average the number of approximates in the region A over a range of heights and similarly average the total number of approximates, we have an almost everywhere equidistribution.

On the other hand, there are examples of lattices Λ and directions $\mathbf{v}$ for which (non-averaged) equidistribution does not hold. The following is proved in [6].

Theorem 2.4 *Let $d \geq 1$. There exists a lattice $\Lambda \in \mathrm{SL}_{d+1}(\mathbb{R})/\mathrm{SL}_{d+1}(\mathbb{Z})$, a set $A \subset \mathbb{S}^{d-1}$ with zero measure boundary, and a sequence $\{T_n\}$ for which*

$$\lim_{n\to\infty} \frac{N(\Lambda, A, \epsilon, T_n)}{N(\Lambda, \epsilon, T_n)} \neq \text{vol}(A) \tag{17}$$

for every $1 > \epsilon \geq 0$.

For $d = 1$, note that $\mathbb{S}^0$:=$\{-1, 1\}$ and we define $\text{vol}(\{-1\}) = \text{vol}(\{1\}) = 1/2$.

In [54], the study of weighted spiraling was taken up and several interesting spiraling and equidistribution results were obtained. Subsequently, in [1], a study of spiraling and equidistribution in number fields was undertaken. Finally, we note that the paper [11] has interesting results on the distribution of approximates for badly approximable numbers.

3 Diophantine Approximation in Number Fields

In this section, we describe some recent advances in Diophantine approximation in number fields. Let k be a number field of degree d over $\mathbb{Q}$, O_k its ring of integers, and S be the set of field embeddings $\sigma : k \hookrightarrow \mathbb{R}$. Then we have $|S| = d$. We will be interested in Diophantine approximation by rationals from a *fixed* number field.

Analogues of Dirichlet's theorem in this setting have been established by several authors (cf. [15, 63, 66]) using appropriate adaptations of the geometry of numbers. Here is Proposition 2.1 from [51], proved using a result from [15], here we are approximating $\mathbf{x} \in \mathbb{R}^d$ by rationals in k.

Proposition 3.1 *There exists a constant and for every* $Q > 0$, *there exists* $p \in O_k$, $q \in O_k \backslash \{0\}$ *with*

$$\|\sigma(q) \cdot \mathbf{x} + \sigma(p)\| \leq C Q^{-1} \text{ and } \|\sigma(q)\| \leq Q.$$

The corollary of Dirichlet's theorem also holds for number fields. Here is Theorem 2.3 in [51].

Theorem 3.2 *There is a constant* $C = C_k > 0$ *depending only on k, such that for every* $x \in k_S$, *there are infinitely many* $p, q \in O_k$ *with* $q \neq 0$ *satisfying:*

$$\|\sigma(q) \cdot \mathbf{x} + \sigma(p)\| \leq C \|\sigma(q)\|^{-1}.$$

We now want to define and discuss the properties of badly approximable vectors. We define the more general *weighted* badly approximable vectors and describe results in [2], we therefore follow the notation from that paper.

Notation 3.3 Let $\mathbf{r} \in \mathbb{R}^d$ be a real vector with $r_\sigma \geq 0$ for $\sigma \in S$ and $\sum_{\sigma\in S} r_\sigma = 1$. Set

$$S_1 = \{\sigma \in S : r_\sigma > 0\}, \text{ and } S_2 = S \setminus S_1.$$

Assume $|S_1| = d_1$, $|S_2| = d_2$. Choose and fix $\omega \in S$ with $r_\omega = r$, where

$$r \;=\; \max_{\sigma\in S} r_\sigma.$$

Definition 3.4 Define a weighted norm, called the $\mathbf{r}$-*norm*, on $\prod_{\sigma\in S}\mathbb{R}$ by

$$\|\mathbf{x}\|_{\mathbf{r}} \;=\; \max_{\sigma\in S_1} |x_\sigma|^{\frac{1}{r_\sigma}}.$$

Definition 3.5 Say a vector $\mathbf{x} = (x_\sigma)_{\sigma\in S} \in \prod_{\sigma\in S}\mathbb{R}$ is $(k, \mathbf{r})$-*badly approximable* if

$$\inf_{q\in O_k\setminus\{0\}} \max_{\mathbb{P}\in O_k} \left\{ \max_{\sigma\in S_1} \|q\|_{\mathbf{r}}^{r_\sigma} |\sigma(q)x_\sigma + \sigma(p)|, \right.$$
$$\left. \max_{\sigma\in S_2} \max\{|\sigma(q)x_\sigma + \sigma(p)|, |\sigma(q)|\} \right\} > 0.$$

The set of $(k, \mathbf{r})$-badly approximable vectors is denoted as $\mathbf{Bad}(k, \mathbf{r})$.

The definition of $(k, \mathbf{r})$-badly approximable vector is the weighted case of k-badly approximable vector introduced in [25]. Weighted badly approximable (by rational) vectors in $\mathbb{R}^n$ are the subject of Schmidt's conjecture, now a theorem of Badziahin, Pollington and Velani [8]. In [2], a number field analogue of Schmidt's conjecture is proved, this was previously known in some special cases [25]. The existence of k-badly approximable vectors was established in [13, 41].

We note in passing that a real number is badly approximable if and only if its partial fraction coefficients are bounded. In [44], this characterization is established for complex numbers and in [43], examples of badly approximable vectors in the number field setting have been constructed.

A variant of Schmidt's game, called the hyperplane potential game was introduced in [28] and defines a class of subsets of $\mathbb{R}^d$ called *hyperplane potential winning* (*HPW* for short) sets.

The hyperplane potential game involves two parameters $\beta \in (0, 1)$ and $\gamma > 0$. Bob starts the game by choosing a closed ball $B_0 \subset \mathbb{R}^d$ of radius ρ_0. In the ith turn, Bob chooses a closed ball B_i of radius ρ_i, and then Alice chooses a countable family of hyperplane neighborhoods

$$\left\{L_{i,k}^{(\delta_{i,k})} \;:\; k \in \mathbb{N}\right\} \;\text{ such that }\; \sum_{k=1}^{\infty} \delta_{i,k}^{\gamma} \;\leq\; (\beta\rho_i)^{\gamma}.$$

Then in the $(i + 1)$th turn, Bob chooses a closed ball $B_{i+1} \subset B_i$ of radius $\rho_{i+1} \geq \beta\rho_i$. By this process, there is a nested sequence of closed balls

$$B_0 \supseteq B_1 \supseteq B_2 \supseteq \cdots.$$

We say a subset $S \subset \mathbb{R}^d$ is (β, γ)-*hyperplane potential winning* ((β, γ)-*HPW* for short) if no matter how Bob plays, Alice can ensure that

$$\bigcap_{i=0}^{\infty} B_i \cap \left(S \cup \bigcup_{i=0}^{\infty} \bigcup_{k=1}^{\infty} L_{i,k}^{(\delta_{i,k})} \right) \neq \varnothing.$$

We say S is *hyperplane potential winning* (HPW) if it is (β, γ)-HPW for any $\beta \in (0, 1)$ and $\gamma > 0$.

Set

$$\theta : k \longrightarrow \prod_{\sigma \in S} \mathbb{R}, \quad \theta(p) = (\sigma(p))_{\sigma \in S}.$$

Let $\mathrm{Res}_{k/\mathbb{Q}}$ denote Weil's restriction of scalar's functor. It is well known that the group $\mathrm{Res}_{k/\mathbb{Q}} \mathrm{SL}_2(\mathbb{Z})$ is a lattice in $\mathrm{Res}_{k/\mathbb{Q}} \mathrm{SL}_2(\mathbb{R})$. The latter coincides with the product of d copies of $\mathrm{SL}_2(\mathbb{R})$. We set

$$G = \mathrm{Res}_{k/\mathbb{Q}} \mathrm{SL}_2(\mathbb{R}) = \prod_{\sigma \in S} \mathrm{SL}_2(\mathbb{R}), \quad \Gamma = \mathrm{Res}_{k/\mathbb{Q}} \mathrm{SL}_2(\mathbb{Z}).$$

It follows from the definition that the subgroup $\mathrm{Res}_{k/\mathbb{Q}} \mathrm{SL}_2(\mathbb{Z})$ coincides with the subgroup $\theta(\mathrm{SL}_2(O_k))$, where θ is the map defined by $\theta(g) = (\sigma(g))_{\sigma \in S}$. The following is a special case of the main theorem in [2].

Proposition 3.6 *Let* $\mathbf{r} \in \mathbb{R}^d$ *be a real vector with* $r_\sigma \geq 0$ *for* $\sigma \in S$ *and* $\sum_{\sigma \in S} r_\sigma = 1$, *set*

$$g_{\mathbf{r}}(t) := \left(\begin{pmatrix} e^{r_\sigma t} & 0 \\ 0 & e^{-r_\sigma t} \end{pmatrix} \right)_{\sigma \in S} \tag{18}$$

and $F_{\mathbf{r}}^+ = \{g_{\mathbf{r}}(t) : t \geq 0\}$, *then the set*

$$E(F_{\mathbf{r}}^+) := \{x \in G/\Gamma : F_{\mathbf{r}}^+ x \text{ is bounded}\}$$

is HPW.

As before, $\mathbf{Bad}(k, \mathbf{r})$ corresponds to bounded orbits for certain flows on homogeneous spaces. Namely, we have the following correspondence (Proposition 3.4 in [2]) between $(k, \mathbf{r})$-badly approximable vector and bounded $F_{\mathbf{r}}^+$ trajectories, i.e. a number field version of Dani's correspondence. The proof is more involved than the $\mathbb{Q}$ case.

Proposition 3.7 *A vector* $\mathbf{x} = (x_\sigma)_{\sigma \in S}$ *is* $(k, \mathbf{r})$-*badly approximable if and only if the trajectory* $F_{\mathbf{r}}^+ u(\mathbf{x})\Gamma$ *is bounded in* G/Γ. *In other words,*

$$\mathbf{Bad}(K, \mathbf{r}) = u^{-1}\left(\pi^{-1}(E(F_{\mathbf{r}}^+)) \cap H\right), \tag{19}$$

where π *denotes the projection* $G \longrightarrow G/\Gamma$.

In view of the number field Dani correspondence above, this implies a number field version of Schmidt's conjecture, in other words.

Theorem 3.8 $\mathbf{Bad}(k, \mathbf{r})$ *is HPW.*

In fact, the sets above are winning for the hyperplane absolute game introduced in [14]. This is because it is proved in [28, Theorem C.8] that A subset S of $\mathbb{R}^d$ is HPW if and only if it is HAW.

4 A Projective Duffin Schaeffer Theorem

In this section, we describe a recent *projective* variation of metric Diophantine approximation originally studied by Choi and Vaaler [18] and developed further by the author and Haynes [31]. Projective metric Diophantine approximation aims to quantify the density of $\mathbb{P}^{n-1}(k)$ in $\mathbb{P}^{n-1}(k_v)$ where k is a number field and k_v is a completion of k.

Notation 4.1 For non-zero vectors $\mathbf{x}, \mathbf{y} \in k_v^n$, we define following [18]

$$\delta_v(\mathbf{x}, \mathbf{y}) := \frac{|\mathbf{x} \wedge \mathbf{y}|_v}{|\mathbf{x}|_v |\mathbf{y}|_v}. \tag{20}$$

Then δ_v defines a metric on $\mathbb{P}^{n-1}(k_v)$ which induces the usual quotient topology [64].

Notation 4.2 We define the height of a point $\mathbf{x} \in \mathbb{P}^{n-1}(k)$ by

$$\mathrm{H}(\mathbf{x}) := \prod_v |\mathbf{x}|_v. \tag{21}$$

We note that this is well defined over projective space because of the product formula. The following is a projective version of Dirichlet's theorem due to Choi and Vaaler [18].

Theorem 4.3 *Let* $\mathbf{x} \in \mathbb{P}^{n-1}(k_v)$*, let* $\tau \in k_v$ *with* $|\tau|_v \geq 1$*. Then there exists* $\mathbf{y} \in \mathbb{P}^{n-1}(k)$ *such that*

1. $\mathrm{H}(\mathbf{y}) \leq c_k(n)|\tau|_v^{n-1}$*, and*
2. $\delta_v(\mathbf{x}, \mathbf{y}) \leq c_k(n)(|\tau|_v \mathrm{H}(\mathbf{y}))^{-1}$.

Here

$$c_k(n) = 2|\Delta_k|^{1/2d} \prod_{v|\infty} r_v(n)^{d_v/d},$$

Δ_k is the discriminant of k, and

$$r_v(n) \;=\; \begin{cases} \pi^{-\frac{1}{2}}\Gamma(\frac{n}{2}+1)^{\frac{1}{n}} & \text{if } v \text{ is real,} \\ (2\pi)^{-\frac{1}{2}}\Gamma(n+1)^{\frac{1}{2n}} & \text{if } v \text{ is complex.} \end{cases}$$

In [31], a projective analogue of Khintchine's theorem and more generally, the Duffin Schaeffer conjecture, were proved. In order to state the results in loc. cit. we first briefly recall some probability measures on $\mathbb{P}^{n-1}(k_v)$, originally defined and studied by Choi [17]. If v is an infinite place then β_v^n is the usual n-fold Lebesgue measure on $\mathbb{R}^n$ or 2^n times Lebesgue measure on $\mathbb{C}^n$, while if v is a finite place then β_v^n is the n-fold Haar measure normalized so that

$$\beta_v(O_v) \;=\; \|\mathcal{D}_v\|_v^{d_v/2},$$

where O_v is the ring of integers of k_v and $\mathcal{D}_v$ is the local different of k at v. Let $\phi : k_v^n \setminus \{\mathbf{0}\} \longrightarrow \mathbb{P}^{n-1}(k_v)$ be the quotient map and define the σ-algebra $\mathcal{M}$ of measurable sets in $\mathbb{P}^{n-1}(k_v)$ to be the collection of sets $M \subseteq \mathbb{P}^{n-1}(k_v)$ such that $\phi^{-1}(M)$ lies in the σ-algebra of Borel sets in k_v^n. This is in fact the σ-algebra of Borel sets in $\mathbb{P}^{n-1}(k_v)$. One then defines measures μ_v on $(\mathbb{P}^{n-1}(k_v), \mathcal{M})$ by

$$\mu_v(M) \;=\; \frac{\beta_v^n\left(\phi^{-1}(M) \cap B(\mathbf{0}, 1)\right)}{\beta_v^n\left(B(\mathbf{0}, 1)\right)}. \tag{22}$$

Given $\psi : \mathbb{R}_+ \cup \{0\} \longrightarrow \mathbb{R}_+ \cup \{0\}$, let $\mathcal{W}$ be the set of $\mathbf{x} \in \mathbb{P}^{n-1}(\mathbb{Q}_v)$ for which there exist infinitely many $\mathbf{y} \in \mathbb{P}^{n-1}(\mathbb{Q})$ such that

$$\delta_v(\mathbf{x}, \mathbf{y}) \;\leq\; \psi(\mathrm{H}(\mathbf{y})).$$

Then it is a straightforward consequence of the Borel–Cantelli lemma that $\mu_p(\mathcal{W}) = 0$ whenever

$$\sum_{q=1}^{\infty} q^{n-1}\psi(q)^{(n-1)} \tag{23}$$

converges. In [31], the projective p-adic version of the Duffin–Schaeffer conjecture is established in all dimensions greater than 1.

Theorem 4.4 *Assume that p is a finite place, that $n > 2$, and that $\psi(q) = 0$ whenever $p|q$. Then $\mu_p(\mathcal{W}_p(\psi, \mathbb{Q}, n)) = 1$ whenever (23) diverges.*

In fact, more can be proved if monotonicity is assumed. The second result in [31] is the complete (i.e., allowing arbitrary primes and dimensions) projective version of Khintchine's theorem.

Theorem 4.5 *Assume that ψ is decreasing and let p be a (finite or infinite) place of $\mathbb{Q}$. Then $\mu_p(\mathcal{W}_p(\psi, \mathbb{Q}, n)) = 1$ whenever (23) diverges.*

Recently, in [40], the authors have studied badly approximable vectors in the projective setting. In particular, they showed that badly approximable vectors have full Hausdorff dimension.

5 The Hyperbolic Picture

Consider the action of $\mathrm{SL}_2(\mathbb{R})$ on the hyperbolic upper half plane $\mathbb{H}^2$ by Möbius transformations. This action extends to the boundary and there are close connections between Diophantine approximation and the study of dense orbits of discrete subgroups of $\mathrm{SL}_2(\mathbb{R})$ on the boundary. For example, the orbit of ∞ under the action of $\mathrm{SL}_2(\mathbb{Z})$ is precisely the set of rational numbers; one might therefore seek a more general quantitative theory of the approximation of limit points of a fixed Kleinian group by points in the orbit (under the group) of a distinguished limit point y. In this section, we briefly review important work by Patterson [62] and then discuss some recent work in this direction carried out in [10].

Let G denote[1] a nonelementary, geometrically finite Kleinian group acting on the unit ball model (B^{d+1}, ρ) of $(d+1)$-dimensional hyperbolic space with metric ρ derived from the differential $d\rho = 2|d\mathbf{x}|/(1-|\mathbf{x}|^2)$. Thus, G is a discrete subgroup of $\mathrm{Möb}(B^{d+1})$, the group of orientation-preserving Möbius transformations of the unit ball B^{d+1}. Since G is nonelementary, the limit set Λ of G is uncountable. The group G is said to be of the first kind (such a group is a lattice) if $\Lambda = S^d$ and of the second kind otherwise. Let δ denote the Hausdorff dimension of Λ. It is well known that δ is equal to the exponent of convergence of the group. For $g \in G$ set $L_g := |g'(0)|^{-1}$, where $|g'(0)| = 1 - |g(0)|^2$ is the (Euclidean) conformal dilation of g at the origin. It can be checked that $L_g \le e^{\rho(0,g(0))} \le 4L_g$. The following two Dirichlet-type theorems were first established by Patterson [62, Sect. 7: Theorems 1 and 2] for finitely generated Fuchsian groups, but can be generalized to higher dimensions.

Theorem 5.1 *Let G be a nonelementary, geometrically finite Kleinian group containing parabolic elements and let P be a complete set of inequivalent parabolic fixed points of G. Then there is a constant $c > 0$ with the following property: for each $\xi \in \Lambda$, $N > 1$, there exist $p \in P$, $g \in G$ so that*

$$|\xi - g(p)| \le \frac{c}{\sqrt{L_g N}} \quad \text{and} \quad L_g \le N.$$

Theorem 5.2 *Let G be a nonelementary, geometrically finite Kleinian group without parabolic elements and let $\{\eta, \eta'\}$ be the pair of fixed points of a hyperbolic element of G. Then there is a constant $c > 0$ with the following property: for all $\xi \in \Lambda$,*

[1] This notation, mainstream in the Kleinian groups literature, is at odds with the notation in previous sections where G was the ambient Lie group and Γ a lattice in G.

$N > 1$, there exist $y \in \{\eta, \eta'\}$, $g \in G$ so that

$$|\xi - g(y)| \leq \frac{c}{N} \quad \text{and} \quad L_g \leq N.$$

As mentioned earlier, in the context of $\mathrm{SL}_2(\mathbb{Z})$ and $\mathbb{H}^2$, Theorem 5.1 reduces to Dirichlet's Theorem.

In [10], the notion of singular limit points within the hyperbolic space setup was introduced. Let G be a Kleinian group and let Y be a complete set P of inequivalent parabolic fixed points of G if the group has parabolic elements; otherwise let Y be the pair $\{\eta, \eta'\}$ of fixed points of a hyperbolic element of G.

Definition 5.3 A point $\xi \in \Lambda$ is said to be *singular* if for every $\varepsilon > 0$ there exists N_0 with the following property: for each $N \geq N_0$, there exist $y \in Y$, $g \in G$ so that

$$|\xi - g(y)| < \begin{cases} \frac{\varepsilon}{\sqrt{L_g N}} & \text{if } Y = P \\ \frac{\varepsilon}{N} & \text{if } Y = \{\eta, \eta'\} \end{cases} \quad \text{and} \quad L_g < N. \tag{24}$$

In [10], it was shown that the hyperbolic singular theory is, irrespective of the dimension of the hyperbolic space, similar to the one-dimensional classical theory.

Theorem 5.4 *Let G be a nonelementary, geometrically finite Kleinian group, and let Y be as above. Then a point $\xi \in \Lambda$ is singular if and only if $\xi \in G(Y) := \{g(y) : g \in G,\ y \in Y\}$.*

In [62], convergence and divergence Khintchine type theorems were proved. In [10], versions of Khintchine's theorem for *proper* subsets of the limit set were investigated. Let K be a subset of the limit set Λ which supports a nonatomic probability measure μ. Given $\alpha > 0$, the measure μ supported on K is said to be *weakly absolutely α-decaying* if there exist strictly positive constants C, r_0 such that for all $\varepsilon > 0$ we have

$$\mu\left(B(x, \varepsilon r)\right) \leq C\, \varepsilon^{\alpha}\, \mu\left(B(x, r)\right) \quad \forall x \in K \quad \forall r < r_0.$$

For sets supporting such measures, the following result was proved in [10].

Theorem 5.5 *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Fix $\alpha > 0$, and let K be a compact subset of Λ equipped with a weakly absolutely α-decaying measure μ. Then*

$$\mu\left(K \cap W_y(\psi)\right) = 0 \quad \text{if} \quad \sum_{r=1}^{\infty} r^{\alpha-1} \psi(r)^{\alpha} < \infty. \tag{25}$$

We now discuss the analogue of badly approximable vectors. The set

$$\mathbf{Bad}_y \;:=\; \left\{\xi \in \Lambda \;:\; \exists c(\xi) > 0 \text{ such that } |\xi - g(y)| > c(\xi)/L_g \;\; \forall g \in G\right\},$$

can be considered to be the hyperbolic analogue of badly approximable numbers and is of measure zero. Nevertheless, it is a large set.

The following theorem was first established by Patterson [62, Sect. 10] for finitely generated Fuchsian groups of the first kind. As before, y is taken to be a parabolic fixed point of G if the group has parabolic elements and a hyperbolic fixed point of G otherwise.

Theorem 5.6 *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Then*

$$\dim \mathbf{Bad}_y \;=\; \dim \Lambda.$$

Let K be a subset of the limit set Λ which supports a nonatomic probability measure μ as before. We assume that the measure μ supported on K is *Ahlfors δ-regular* for some $\delta > 0$; that is, that there exist constants $C > 0$ and r_0 such that

$$C^{-1} r^{\delta} \;\leq\; \mu\left(B(x,r)\right) \;\leq\; C\, r^{\delta} \;\; \forall x \in K \;\; \forall r < r_0.$$

Sets supporting such measures are referred to as Ahlfors δ-regular and it is a well known fact that

$$\dim K = \delta.$$

For Ahlfors δ-regular subsets of the limit set the following result was proved in [10].

Theorem 5.7 *Let G be a nonelementary, geometrically finite Kleinian group and let y be a parabolic fixed point of G, if there are any, and a hyperbolic fixed point otherwise. Let K be a compact, Ahlfors δ-regular subset of Λ. Then*

$$\dim\left(K \cap \mathbf{Bad}_y\right) = \dim K. \tag{26}$$

These results were motivated by results on Diophantine approximation on manifolds, and indeed, constitute an hyperbolic analogue of the theory. We refer to [10] for details.

Acknowledgements This survey grow out of lectures delivered at CIRM Luminy, at the University of Houston and at IIT Delhi. I am grateful to the organisers of each of the three events for inviting me and for their hospitality. Special thanks to Jayadev Athreya, Alan Haynes and Riddhi Shah. I would also like to thank the editors of this volume, Anima Nagar, Riddhi Shah and Shrihari Sridharan. This work was supported by a grant from the Indo-French Centre for the Promotion of Advanced Research; a Department of Science and Technology, Government of India Swarnajayanti fellowship; a MATRICS grant from the Science and Engineering Research Board; and the Benoziyo Endowment Fund for the Advancement of Science at the Weizmann Institute. I gratefully acknowledge the hospitality of the Technion and the Weizmann Institute.

References

1. Alam, M., & Ghosh, A. (2020). Equidistribution on homogeneous spaces and the distribution of approximates in Diophantine approximation. *Transactions of the American Mathematical Society, 373*(5), 3357–3374.
2. An, J., Ghosh, A., Guan, L., & Ly, T. (2019). Bounded orbits of diagonalizable flows on finite volume quotients of products of SL(2, $\mathbb{R}$). *Advances in Mathematics,354,* 106743, 18 pp.
3. Athreya, J. S., & Ghosh, A. (2018). The Erdős-Szüsz-Turán distribution for equivariant processes. *L'Enseignement Mathématique, 64*(2), 1–21.
4. Athreya, J. S., Ghosh, A., & Prasad, A. (2012). Ultrametric logarithm laws II. *Monatshefte für Mathematik, 167*(3), 333–356.
5. Athreya, J. S., Ghosh, A., & Tseng, J. (2014). Spherical averages of Siegel transforms for higher rank diagonal actions and applications. arXiv:1407.3573.
6. Athreya, J. S., Ghosh, A., & Tseng, J. (2015). Spherical averages of Siegel transforms and spiraling of lattice approximations. *Journal of the London Mathematical Society, 91*(2), 383–404.
7. Athreya, J. S., Biswas, K., & Ghosh, A. (2016). Diophantine approximation and lattice actions on the Clifford plane. In L. Ji, A. Papadopoulos & W. Su (Eds.), *Teichmüller theory and its impact*.
8. Badziahin, D., Pollington, A., & Velani, S. (2011). On a problem in simultaneous Diophantine approximation: Schmidt's conjecture. *Annals of Mathematics, 174*(3), 1837–1883.
9. Beresnevich, V., Ramírez, F., & Velani, S. (2016). Metric Diophantine approximation: Aspects of recent work. *Dynamics and analytic number theory*. London mathematical society lecture note series (Vol. 437, pp. 1–95). Cambridge: Cambridge University Press.
10. Beresnevich, V., Ghosh, A., Simmons, D., & Velani, S. (2018). Diophantine approximation in Kleinian groups: Singular, extremal, and bad limit points. *Journal of the London Mathematical Society, 98*(2), 306–328.
11. Bilyk, D., Ma, X., Pipher, J., & Spencer, C. (2016). Diophantine approximation and directional discrepancy of rotated lattices. *Transactions of AMS, 368,* 3871–3897.
12. Borel, A. (1995). Values of indefinite quadratic forms at integral points and flows on spaces of lattices. *Bulletin of the American Mathematical Society, 32,* 184–204.
13. Bourgain, J. (2016). A quantitative Oppenheim theorem for generic diagonal quadratic forms. *Israel Journal of Mathematics, 215*(1), 503–512.
14. Broderick, R., Fishman, L., Kleinbock, D., Reich, A., & Weiss, B. (2012). The set of badly approximable vectors is strongly C^1 incompressible. *Mathematical Proceedings of the Cambridge Philosophical Society,153*(2), 319–339.
15. Burger, E. (1992). Homogeneous Diophantine approximation in S-integers. *Pacific Journal of Mathematics, 152*(2), 211–253.
16. Cassels, J. (1957). *An introduction to Diophantine approximation*. Cambridge tracts in mathematics and mathematical physics (Vol. 45). Cambridge: Cambridge University Press.
17. Choi, K. (2000). On the distribution of points in projective space of bounded height. *Transactions of the American Mathematical Society, 352*(3), 1071–1111.
18. Choi, K., & Vaaler, J. D. (1999). Diophantine approximation in projective space. *Number theory (Ottawa, ON, 1996)*. CRM proceedings and lecture notes (Vol. 19, pp. 55–65). Providence: American Mathematical Society.
19. Chow, S., Ghosh, A., Guan, L., Marnat, A., & Simmons, D. (2018). Diophantine transference inequalities: Weighted, inhomogeneous, and intermediate exponents. *Annali della Scuola Normale Superiore di Pisa*. arXiv:1808.07184.
20. Dani, S. G. (1985). Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *Journal für die Reine und Angewandte Mathematik, 359,* 55–89.
21. Dani, S. G. (2020). Article in this volume.
22. Davenport, H., & Schmidt, W. M. (1970). Dirichlet's theorem on Diophantine approximation: II. *Acta Arithmetica, 16,* 413–424.

23. Einsiedler, M., & Ward, T. (2011). *Ergodic theory with a view towards number theory*. Graduate texts in mathematics (Vol. 259, pp. xviii+481). London: Springer.
24. Einsiedler, M., & Ward, T. (2016). Diophantine problems and homogeneous dynamics. *Dynamics and analytic number theory*. London mathematical society lecture note series (Vol. 437, pp. 258–288). Cambridge: Cambridge University Press.
25. Einsiedler, M., Ghosh, A., & Lytle, B. (2016). Badly approximable vectors, C1 curves and number fields. *Ergodic Theory and Dynamical Systems, 36*(6), 1851–1864.
26. Erdős, P., Szüsz, O., & Turán, P. (1958). Remarks on the theory of Diophantine approximation. *Colloquium Mathematicum, 6,* 119–126.
27. Eskin, A. (2010). Unipotent flows and applications. *Homogeneous flows, moduli spaces and arithmetic*. Clay mathematics proceedings (Vol. 10, pp. 71–129). Providence: American Mathematical Society.
28. Fishman, L., Simmons, D., & Urbański, M. (2018). Diophantine approximation and the geometry of limit sets in Gromov hyperbolic metric spaces. *Memoirs of the American Mathematical Society,254*(1215), v+137.
29. Ganguly, A., & Ghosh, A. (2017). Dirichlet's theorem in function fields. *Canadian Journal of Mathematics, 69,* 532–547.
30. Ghosh, A. (2018). Diophantine approximation on subspaces of n and dynamics on homogeneous spaces. *Handbook of group actions. Vol. IV*. Advanced lectures in mathematics (ALM) (Vol. 41, pp. 509–527). Somerville: International Press.
31. Ghosh, A., & Haynes, A. (2016). Projective metric number theory. *Journal für die Reine und Angewandte Mathematik, 712,* 39–50.
32. Ghosh, A., & Kelmer, D. (2017). Shrinking targets for semisimple groups. *Bulletin of the London Mathematical Society, 49*(2), 235–245.
33. Ghosh, A., & Kelmer, D. (2018). A quantitative Oppenheim theorem for generic ternary quadratic forms. *Journal of Modern Dynamics, 12,* 1–8.
34. Ghosh, A., & Royals, R. (2015). An extension of the Khintchine-Groshev theorem. *Acta Arithmetica, 167,* 1–17.
35. Ghosh, A., Gorodnik, A., & Nevo, A. (2013). Diophantine approximation and automorphic spectrum. *International Mathematics Research Notices, 21,* 5002–5058.
36. Ghosh, A., Gorodnik, A., & Nevo, A. (2014). Metric Diophantine approximation on homogeneous varieties. *Compositio Mathematica, 150*(8), 1435–1456.
37. Ghosh, A., Gorodnik, A., & Nevo, A. (2015). Diophantine approximation exponents on homogeneous varieties. *Contemporary Mathematics, 631,* 181–200.
38. Ghosh, A., Gorodnik, A., & Nevo, A. (2018). Best possible rates of distribution of dense lattice orbits in homogeneous spaces. *Journal für die Reine und Angewandte Mathematik, 745,* 155–188.
39. Ghosh, A., Gorodnik, A., & Nevo, A. (2018). Optimal density for values of generic polynomial maps. *American Journal of Mathematics*. arXiv:1801.01027.
40. Harrap, S., & Hussain, M. (2017). A note on badly approximable sets in projective space. *Mathematische Zeitschrift, 285,* 239–250.
41. Hattori, T. (2007). Some Diophantine approximation inequalities and products of hyperbolic spaces. *Journal of the Mathematical Society of Japan, 59,* 239–264.
42. Hill, R., & Velani, S. (1995). The ergodic theory of shrinking targets. *Inventiones Mathematicae, 119,* 175–198.
43. Hines, R. (2017). Badly approximable numbers over imaginary quadratic fields. arXiv:1707.07231.
44. Hines, R. (2018). Examples of badly approximable vectors over number fields. arXiv:1809.07404.
45. Jarnik, V. (1929). Diophantischen approximationen und Hausdorffsches mass. *Matematicheskii Sbornik, 36,* 371–382.
46. Kesten, H. (1962). Some probabilistic theorems on Diophantine approximations. *Transactions of the American Mathematical Society, 103,* 189–217.

47. Kesten, H., & Sós, V. (1966). On two problems of Erdös, Szüsz and Turán concerning Diophantine approximations. *Acta Arithmetica, 12,* 183–192.
48. Khintchine, A. (1926). Über eine klasse linearer Diophantische approximationen. *Rendiconti del Circolo Matematico di Palermo, 50,* 170–195.
49. Kleinbock, D. (2001). Some applications of homogeneous dynamics to number theory. *Smooth ergodic theory and its applications (Seattle, WA, 1999)*. Proceedings of symposia in pure mathematics (Vol. 69, pp. 639–660). Providence: American Mathematical Society.
50. Kleinbock, D. (2010). Quantitative nondivergence and its Diophantine applications. *Homogeneous flows, moduli spaces and arithmetic*. Clay mathematics proceedings (Vol. 10, pp. 131–153). Providence: American Mathematical Society.
51. Kleinbock, D., & Ly, T. (2016). Badly approximable S-numbers and absolute Schmidt games. *Journal of Number Theory,164,* 13–42.
52. Kleinbock, D., & Margulis, G. A. (1999). Logarithm laws for flows on homogeneous spaces. *Inventiones Mathematicae, 138,* 451–494.
53. Kleinbock, D., Shah, N., & Starkov, A. (2002). Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory. *Handbook of dynamical systems* (Vol. 1A, pp. 813–930). Amsterdam: North-Holland.
54. Kleinbock, D., Shi, R., & Weiss, B. (2017). Pointwise equidistribution with an error rate and with respect to unbounded functions. *Mathematische Annalen, 367*(1–2), 857–879.
55. Mahler, K. (1961). *Lectures on Diophantine approximations* (pp. 181–188). Notre Dame: University of Notre Dame.
56. Margulis, G. (1987). Formes quadratriques indéfinies et flots unipotents sur les espaces homogènes. *Comptes Rendus de l'Académie des Sciences Paris Séries I - Mathematics,304*(10), 249–253.
57. Marklof, J. (2000). The n-point correlations between values of a linear form. *Ergodic Theory and Dynamical Systems,20*(4), 1127–1172.
58. Maucourant, F. (2009). Arithmetical and geometrical aspects of homogeneous Diophantine approximation by algebraic numbers in a given number field. *Dynamical systems and Diophantine approximation*. Séminaires and congres (Vol. 19, pp. 37–48). Paris: Société Mathématique de France.
59. Morris, D. W. (2005). *Ratner's theorems on unipotent flows*. Chicago lectures in mathematics. Chicago: University of Chicago Press.
60. Oh, H. (2010). Orbital counting via mixing and unipotent flows. *Homogeneous flows, moduli spaces and arithmetic*. Clay mathematics proceedings (Vol. 10, pp. 339–375). Providence: American Mathematical Society.
61. Patterson, S. J. (1976). The limit set of a Fuchsian group. *Acta Mathematica, 136,* 241–273.
62. Patterson, S. J. (1976). Diophantine approximation in Fuchsian groups. *Philosophical Transactions of the Royal Society London Series A, 282,* 527–563.
63. Quéme, R. (1992). On Diophantine approximation by algebraic numbers of a given number field: A new generalization of Dirichlet approximation theorem. *Journées Arithmétiques* (1989) (Luminy, 1989), *Astérisque,198–200* (1991), 273–283.
64. Rumely, R. S. (1989). *Capacity theory on algebraic curves*. Lecture notes in mathematics (Vol. 1378). New York: Springer.
65. Schmidt, W. M. (1966). On badly approximable numbers and certain games. *Transactions of the American Mathematical Society, 123,* 178–199.
66. Schmidt, W. M. (1975). Simultaneous approximation to algebraic numbers by elements of a number field. *Monatshefte für Mathematik, 79,* 55–66.
67. Schmidt, W. M. (1980). *Diophantine approximation*. Lecture notes in mathematics (Vol. 785). Berlin: Springer.
68. Sullivan, D. (1983). Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta Mathematica, 144,* 215–237.

On Certain Unusual Large Subsets Arising as Winning Sets of Some Games

S. G. Dani

1 Introduction

Consider the space $\mathbb{R}$ of real numbers. When would we call a subset X of $\mathbb{R}$ a large set? Of course, the whole of $\mathbb{R}$ itself or a subset missing only finitely many points would readily qualify to be large. With some understanding of cardinals, we may add to this list the class of subsets whose complements are countable. This includes for instance the sets of all irrational numbers, the set of all transcendental numbers etc. and we recognize these as large sets.

In topology, we encounter also another kind of sets which are considered large, viz., intersections of countably many open dense sets or, equivalently, sets whose complement is a countable union of closed sets with no interior point (nowhere dense closed sets). In a general topological space, this condition would not be adequate to ensure largeness; for instance, in the space of rational numbers every subset, including the empty subset meets the condition. However, the Baire category theorem tells us that for many "natural" topological spaces that are of interest to analysts, and dynamicists, viz., when the space is either a complete metric space or a locally compact Hausdorff space, it gives a nice criterion to distinguish a class of subsets as large; in this case the class is precisely the class of of all dense $\mathcal{G}_\delta$ subsets; recall that a subset X of a topological space is said to be $\mathcal{G}_\delta$ if it can be expressed as the intersection of a countable family of dense open subsets of the space. In particular, in the case of $\mathbb{R}$ we have the class of $\mathcal{G}_\delta$ subsets as a collection of large sets.

Measure theory provides yet another class of large subsets of $\mathbb{R}$, and more generally of spaces equipped with a measure, including $\mathbb{R}^n$ for all $n \geq 1$. On $\mathbb{R}$, we consider the Lebesgue measure, and in many contexts a set of measure 0 qualifies to be treated as "negligible" and hence sets may considered large if they are of full measure, namely with the complement having measure 0.

S. G. Dani (✉)
Centre for Excellence in Basic Sciences (CEBS), University of Mumbai, Mumbai, India
e-mail: shrigodani@cbs.ac.in

A. Nagar et al. (eds.), *Elements of Dynamical Systems*, Texts and Readings in Mathematics 79, https://doi.org/10.1007/978-981-16-7962-9_7

Thus, in classical mathematics, there are a variety of subsets that serve as classes of "large" subsets. We note that largeness in the sense of being a $\mathcal{G}_\delta$ is distinct from being large in the measure-theoretic sense as above. There exist $\mathcal{G}_\delta$ sets of measure 0 while on the other hand there are sets whose complement has zero measure, that are not $\mathcal{G}_\delta$. Thus, each of the collections can be thought of as a collection of sets that are large in their own way, in their own context.

It would seem hard to think up of other classes of sets, independently, that ought to be considered large, in some heuristically valid sense. However yet another class of large sets has turned up and is involved any many studies in recent times, in number theory and dynamics. These are classes of sets that are "winning sets" for certain games. It is the aim of this talk to introduce these and the various contexts in which they appear, and some recent results about them.

All games considered will be two player games, with infinitely many turns each. We shall call the players Alice and Bob. Before going over to some of the main games of interest in the overall context, we shall discuss a toy version. I shall call it the *number building game*.

Consider a number written in the decimal representation in the form

$$\alpha \;=\; 0\,.\,a_1\,b_1\,a_2\,b_2\,\ldots,$$

where $a_1,\ a_2,\ \ldots$ and $b_1,\ b_2,\ \ldots$ are digits from $\{0,\ 1,\ \ldots,\ 9\}$; the digits $a_1,\ a_2,\ \ldots$ are picked by Alice and $b_1,\ b_2,\ \ldots$ are picked by Bob taking turns, adding one digit at a time alternately, starting with a_1, then b_1, etc. The (infinite) process of their picking the digits alternately produces a number $\alpha \in [0, 1]$. Now, the objective of the game is as follows.

Let S be a given subset of $[0, 1]$. Alice would be the winner if α belongs to S and Bob would be the winner if α does not belong to S. Can Alice make the choices of $a_1,\ a_2,\ \ldots$ during her turns, in such a way that irrespective of what choices Bob makes for $b_1,\ b_2,\ \ldots$, adversarially, during his turns, so as to ensure the resulting number α to be in S? This would, of course depend on the set S, and if answer is in the affirmative we say that S is a winning set (we are being partial to Alice here in terms of the terminology, by abbreviating "winning set for Alice" as simply "winning set"; we will be indulging in such a partiality in the sequel as well). It is not difficult to see that if S is a set whose complement in $[0, 1]$ is finite, or countable, then S is a winning set, and that winning sets have to be uncountable, while on the other hand the complement of a set of zero measure (necessarily uncountable and large in measure theoretic sense) need not always be a winning set; the set of numbers in which a particular digit, say 5, occurs in the decimal expansion with asymptotic frequency $\frac{1}{10}$ has Lebesgue measure 1 (this is implied by the ergodic theorem, or the strong law of large numbers, but can also be verified by direct computation), but it is not a winning set, since Bob can choose 5 at all turns, in which case the asymptotic frequency, if it exists, would be at least $\frac{1}{2}$.

The game as above can evidently also be considered with respect to any base in place of 10. Also, in place of a_i and b_i being digits we can have them to be blocks of digits. Analogous assertions to the above hold in these cases as well.

Before moving ahead, the reader is alerted that the games considered in the following sections differ from the one above in one respect. While here at each successive stage only finitely many options were involved, in general there will be a continuous family of options possible, for both the players.

2 Schmidt's (α, β)-Game

We next introduce a game which was ushered in by W. M. Schmidt. It has been a prototype for various games studied in literature subsequently and applied in various contexts, involving an idea of large sets; some of these will be discussed in the sequel. Though we shall largely be concerned with situations where the underlying space is the euclidean space $\mathbb{R}^n$ or the torus $\mathbb{T}^n$, for some $n \geq 1$, it would be convenient to introduce the game in the generality of metric spaces.

Let X be a locally complete metric space, viz., every point in X has a neighbourhood which is complete as a metric space.[1] By a ball in X we shall always mean ball of positive radius, and for a ball B the radius will be denoted by $r(B)$. We shall have the players Alice and Bob play a game with closed balls in X as follows. We shall assign numbers $\alpha,\ \beta \in (0, 1)$ to Alice and Bob respectively. The game shall begin by Bob choosing a complete ball in X (viz., complete as a metric space with the induced metric); we denote the ball by B_0; remember we view Bob as the controller of the setting, and Alice as a challenger, one to whose winning our anxieties are linked, so it is proper that Bob gets to make the opening choice; as we shall see, this way of organizing also brings some neatness to the results for us.

The sample play shall continue as follows:

Alice chooses a closed ball of $A_1 \subset B_0$, with $r(A_1) = \alpha r(B_0)$;

Bob chooses a closed ball of $B_1 \subset A_1$, with $r(B_1) = \beta r(A_1) = \alpha\,\beta\, r(B_0)$;

Alice chooses a closed ball of $A_2 \subset B_1$, with $r(A_2) = \alpha r(B_1) = \alpha^2\,\beta\, r(B_0)$;

Bob chooses a closed ball of $B_2 \subset A_2$, with $r(B_2) = \beta r(A_2) = \alpha^2\,\beta^2\, r(B_0)$;

....

Proceeding in this way, we get sequences of closed balls $\{A_i\}$ and $\{B_i\}$ such that $A_i \subset B_{i-1}$ and $B_i \subset A_i$, $r(A_i) = \alpha r(B_{i-1})$ and $r(B_i) = \beta r(A_i)$ for all $i \geq 1$. Then $\{A_i\}$ and $\{B_i\}$ are decreasing sequences of closed balls with diameters tending to 0 and since B_0 is complete it follows that the intersections $\bigcap_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} B_i$ consist of single points, which moreover has to be the same point. Thus, $\bigcap_i A_i = \bigcap_i B_i = \{p\}$, for some $p \in X$.

[1] It has been customary to assume X itself to be complete. However, it would be convenient to have this broader setting, so that the discussion applies also when X is an open subset of a Euclidean space.

Definition 2.1 We say that a subset S of X is (α, β)-*winning* if Alice can ensure, through her choices of A_i, the point of intersection p to be in S. We shall further say that S is α-*winning* if it is (α, β)-winning for all $\beta \in (0, 1)$.

While the whole space is obviously an (α, β)-winning set for any $\alpha, \beta \in (0, 1)$, interestingly, some further conditions need to be met by the pair (α, β) for there to exist proper subsets which are (α, β)-winning; it was noted by Schmidt [9] that for this we need that $1 - 2\alpha + \alpha\beta > 0$. We shall give here a simple proof of this in the case when X is an open subset of $\mathbb{R}^n$, $n \geq 1$; (see [9] for a more general technical version); the metric involved is meant to be the usual one, but the proof works also for various other metrics and, more generally, on Banach spaces, manifolds etc. The crucial property of the metric that is used in the proof is that given two closed balls $B(a, r)$ and $B(b, s)$ the latter is contained in the former if and only if $\|b - a\| \leq r - s$.

Lemma 2.2 *Let X be an open subset of $\mathbb{R}^n$, $n \geq 1$, and suppose that $1 - 2\alpha + \alpha\beta \leq 0$. Then, no proper subset is an (α, β)-winning set in X.*

Proof It suffices to show that for any $p \in X$, $X \backslash \{p\}$ is not an (α, β)-winning set. Let $p \in X$ be given. We shall show that when the condition in the hypothesis holds, there is a strategy by which Bob can ensure that p is contained in B_i for all $i \geq 0$. He will start with B_0 to be a closed ball $B(p, r)$, for some $r > 0$, with centre at p. If a is the centre of the closed ball A_1 picked by Alice, then $A_1 = B(a, \alpha r)$, and since it is contained in $B(p, r)$, by the observation above we have $\|a - p\| \leq (1 - \alpha)r \leq (\alpha - \alpha\beta)r$, by the condition in the hypothesis.

Thus, $\|p - a\| \leq (\alpha - \alpha\beta)r = (1 - \beta)\alpha r$. In turn, the above observation now implies that $B(p, \alpha\beta r)$ is contained in $B(a, \alpha r)$. Thus, Bob can choose B_1 to be $B(p, \alpha\beta r)$. In the same way, for all i he will be able to choose B_i to have its centre at p. Thus $\bigcap_{i=1}^{\infty} B_i = \{p\}$, so $X \backslash \{p\}$ is not an (α, β)-winning set in X. □

The lemma, in particular, shows that a proper subset of $\mathbb{R}^n$ can be α-wining only for $\alpha \leq \frac{1}{2}$.

Remark 2.3 Let $\alpha, \beta \in (0, 1)$ be such that $1 - 2\alpha + \alpha\beta > 0$. A perusal of the proof of Lemma 2.2 shows that, conversely, given any p, Bob can be prevented from choosing B_i to be centered at p, and it could further be ensured at $p \notin B_i$ for some i. This means that in this case for any $\alpha, \beta \in (0, 1)$, $\mathbb{R}^n \backslash \{a\}$ is an (α, β)-winning set. Moreover, given a sequence $\{a_j\}$ it can ensured, sequentially, that each $\{a_j\}$ is outside B_{i_j} for some i_j, and therefore $\mathbb{R}^n \backslash \{a_j\}$ is an (α, β)-winning set for α, β as above. Thus complements of countable subsets are α-winning for all $\alpha \in (0, \frac{1}{2}]$.

Remark 2.4 It is clear that for a set S to be (α, β)-winning it has to be dense in the metric space, since otherwise Bob can choose B_0 itself to be contained outside S. Using the idea as in the proof of Lemma 2.2, it can be seen that if $1 - 2\beta + \alpha\beta \leq 0$, namely if $2\beta \geq 1 + \alpha\beta$, then every dense subset is (α, β)-winning. This makes the case when $2\beta \geq 1 + \alpha\beta$ uninteresting from the point of view of large set theory.

3 Largeness of Winning Sets

Schmidt [9] also proved a host of interesting properties of winning sets, which duly qualify them to be viewed as large sets. Firstly, here is a result about their Hausdorff dimension (see section chapter [4], Sect. (3.2.2), for a precise definition).

Theorem 3.1 ([9], Sect. 11, Corollary 2) *Let $\alpha \in (0, 1)$ and S be an α-winning set in $\mathbb{R}^n$, $n \geq 1$. Then S has Hausdorff dimension n.*

Moreover, it can be seen from the proof that for S as in Theorem 3.1, for any nonempty open subset Ω, the Hausdorff dimension of $S \cap \Omega$ is n; in particular, $S \cap \Omega$ is uncountable. The proof also shows that the conclusion holds for the winning sets in open subsets of $\mathbb{R}^n$ in place of $\mathbb{R}^n$ itself.

It may be noted here that for any (α, β)-winning set the corollary from [9] cited above provides a lower estimate for the Hausdorff dimension of S as a function of α and β, which converges to n as $\beta \to 0$, leading to Theorem 3.1 as above.

Another largeness feature of α-winning sets is that intersection of any two of them, and even countably infinitely many of them, is also α-winning; that is, we have:

Theorem 3.2 ([9], Sect. 6, Theorem 2) *Let $\alpha \in (0, 1)$ and $\{S_j\}$ be a sequence of α-winning sets. Then $\bigcap_{j=1}^{\infty} S_j$ is an α-winning set.*

Proof Let $S = \bigcap_{j=1}^{\infty} S_j$ and consider the (α, β) game for a given $\beta \in (0, 1)$. For $j \geq 1$, let $\beta_j = \beta(\alpha\beta)^{2^j-1}$. A strategy for Alice to ensure the point of intersection to be in S, under the (α, β) game, can then be produced as a combination of winning strategies for S_j to be (α, β_j)-winning, that Alice may fix for each $j = 1, 2, \ldots$. For this, we note that for any $j \in \mathbb{N}$, the sequences of closed balls $\{B_{2i}\}_{i=1}^{\infty}$ and $\{A_{2i+1}\}_{i=0}^{\infty}$ correspond in a natural way to sample sequences in a (α, β_1)-game, with the initial choice of Bob being B_0. More generally, for each $j = 1, 2, \ldots$, the closed ball $A_{2^{j-1}(2i-1)}$ corresponds to the choice at the ith turn in the (α, β_j)-game, with the initial choice of Bob given by $B_{2^{j-1}-1}$. Every natural number k can be realised uniquely as $2^{j-1}(2i + 1)$, with $i, j \in \mathbb{N}$ uniquely defined, and for this k, Alice can choose A_k to be the ball that she would choose at the ith turn according to the strategy selected for S_j to be (α, β_j)-winning. The point of intersection is then assured to be in $\bigcap S_j = S$, as desired. As this holds for all $\beta \in (0, 1)$ this shows that S is an α-winning set. □

We next state a variation of a result of [9], noted earlier in [4] (proposition (5.3), for complete metric spaces); Schmidt considered a class of maps that he called "local isometries" in place of the bi-Lipschitz condition in the statement below. The proof of the assertion is straightforward and will be omitted.

Theorem 3.3 *Let X and Y be complete metric spaces and $f : X \longrightarrow Y$ be a map such that the restriction of f to any bounded subset of X is a bi-Lipschitz map. Let*

S be an α-winning set in Y, where $\alpha \in (0, 1)$. Then $f^{-1}(S)$ is an α-winning subset of X.

In particular, it can be deduced from this that if S is an α-winning set for some $\alpha \in (0, 1)$ and f is a local diffeomorphism of $\mathbb{R}^d$, then $f^{-1}(S)$ is α-winning. Together the theorems imply that if S is an α-winning set in $\mathbb{R}^d$ and f_j is a sequence of local diffeomorphisms of $\mathbb{R}^d$, then $\bigcap_{j=1}^{\infty} f_j(S)$ is an α-winning set, and in particular its intersection with any open subset is of Hausdorff dimension d; we note that d is the maximum possible Hausdorff dimension for a subset of $\mathbb{R}^d$.

4 Large Sets Involved in Diophantine Approximation

In [9], Schmidt established the (α, β)-winning property for all $\alpha, \beta \in (0, 1)$ with $1 - 2\alpha + \alpha\beta > 0$, for the set of what are called badly approximable numbers. We recall that $\alpha \in \mathbb{R}$ is said to be badly approximable if there exists a $\delta > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| > \frac{\delta}{q^2} \ \forall p, q \in \mathbb{N},\ q \neq 0.$$

It may also be recalled that a number is badly approximable if and only if the partial quotients in its continued fraction expansion are bounded.

Theorem 4.1 ([9], Sect. 7, Theorem 3) *The set of badly approximable numbers is (α, β)-winning for all $\alpha, \beta \in (0, 1)$ such that $1 - 2\alpha + \alpha\beta > 0$; in particular it is α-winning for all $\alpha \in (0, \frac{1}{2}]$.*

Proof Let $\gamma = 1 - 2\alpha + \alpha\beta > 0$ and let $k \in \mathbb{N}$ be such that $(\alpha\beta)^k < \gamma$. Let $M = (\alpha\beta)^{-k/2}$. Let B_0 be the initial closed ball, a closed interval in this instance, say of length $2r_0$; without loss of generality we may assume r_0 to be less than $\frac{1}{2}\gamma$ (as Alice can wait to apply the strategy until the radius of the ball chosen by Bob is small enough). Let $\delta = r_0(\gamma - (\alpha\beta)^k)$. We shall show that Alice can play in such a way that if $\{B_i\}$ is the sequence of closed balls chosen by Bob, then for any $x \in B_{kn}$, $n \geq 0$, we have $|x - \frac{p}{q}| \geq \frac{\delta}{q^2}$ for all $p, q \in \mathbb{Z}$, with $0 < q < M^n$. Thus, centred at each rational point p/q, there is an interval of length $2\delta/q^2$ that Alice has to get out of, and for $0 < q < M^n$ this will be achieved at the knth turn. This then readily implies that the point of intersection is a badly approximable number.

The desired statement holds trivially for $n = 0$ (as there is no q satisfying the condition) so it suffices to assume that it holds for $0, \ldots, n-1$, $n \geq 1$, and uphold it for n. Consider the set, say E, of rationals $\frac{p}{q}$ such that $M^{n-1} \leq q < M^n$; we note that if $\frac{p}{q}$ and $\frac{p'}{q'}$ are distinct elements of E then we have

$$\left|\frac{p}{q} - \frac{p'}{q'}\right| \geq \frac{1}{qq'} > M^{-2n}.$$

On the other hand, when $\frac{p}{q} \in E$ we have

$$\frac{\delta}{q^2} \;\leq\; \delta M^{-2(n-1)} \;<\; \frac{1}{4}M^{-2n},$$

by the condition on δ. Thus any two intervals corresponding the elements from E are separated by a distance at least $\frac{1}{2}M^{-2n}$.

The length of the interval $B_{k(n-1)}$ is $2r_0(\alpha\beta)^{k(n-1)}$, which is less than $\frac{1}{2}M^{-2n}$, by the choice of M. Hence $B_{k(n-1)}$ intersects at most one of the intervals corresponding to an element from E. If there is no such interval then we are through, since B_{kn} is contained in $B_{k(n-1)}$. Now suppose there exists an element say $\frac{p}{q}$ of E intersecting $B_{k(n-1)}$. Let $B_{k(n-1)} = [a - r, a + r]$, where $a \in \mathbb{R},\ r > 0$, and suppose for definiteness that $\frac{p}{q} \leq a$; the other case follows in a similar way, symmetrically. Let $A_{k(n-1)+1}$ be chosen to be the interval $[a + (1 - 2\alpha)r, a + r]$, which indeed has the desired radius. The next interval $B_{k(n-1)+1}$ chosen by Bob has the form $[b - \alpha\beta r, b + \alpha\beta r]$, with $b - \alpha\beta r \geq a + (1 - 2\alpha)r$; and hence $b \geq a + \gamma r$. Following the strategy of choosing the rightmost interval of requisite length for the successive $k - 1$ turns, Alice can also ensure that if c is the center of B_{kn} then $c > a + \gamma r$. Now $B_{kn} = [c - (\alpha\beta)^k r, c + (\alpha\beta)^k r]$, so for all $x \in B_{kn}$, we have $x > a$, and hence

$$\left|x - \frac{p}{q}\right| \;\geq\; (x - a) \;\geq\; c - (\alpha\beta)^k r - a \;\geq\; (\gamma - (\alpha\beta)^k)r \;=\; \delta\frac{r}{r_0}.$$

On the other hand,

$$r \;=\; (\alpha\beta)^{k(n-1)} r_0 \;=\; M^{-2(n-1)} r_0 \;\geq\; \frac{r_0}{q^2},$$

which shows that

$$\left|x - \frac{p}{q}\right| \;\geq\; \frac{\delta}{q^2}.$$

This completes the inductive step, and hence the proof of the proposition. □

It would be an interesting exercise for the reader to show, along the lines of the above proof, that the set of badly approximable numbers is also a winning set for the number building game described in (Sect. 1).

Corollary 4.2 *Given a sequence of differentiable functions* $\{f_j\}$ *on* $\mathbb{R}$ *whose derivatives are nowhere-vanishing, the set*

$$\left\{t \in \mathbb{R} \;\mid\; f_j(t) \text{ is badly approximable for all } j\right\}$$

is of Hausdorff dimension 1 *(in particular, it is an uncountable set).*

Analogously to numbers, there are also notions of badly approximable vectors in Euclidean spaces, and badly approximable systems of linear forms, meant to capture

the sense of how well they can be approximated by rational systems of the same kind. Analogous results have been obtained in these cases concerning α-winning nature of the systems that are badly approximable in the respective framework; it must be mentioned that higher dimensional situations involve some additional intricacies. The reader is referred to [5, 9, 10] and other references there for further details.

5 Large Sets in Geometry and Dynamics

We shall now discuss the notion of winning sets in general geometric and dynamical contexts.

5.1 Winning Sets in $\mathbb{R}^d$

The following generalization of Theorem 4.1 was proved in [2], through an analogous proof; apart from extending to higher dimensions the class of sets involved is also more general, even in dimension 1. The generalized version was applied to discuss the class of bounded geodesics on manifolds, and more generally 'orbifolds', of constant negative curvature, which are noncompact but have finite volume; see (Sect. 7.5.3).

Theorem 5.1 *Let* $\{v_i\}$ *be a sequence of vectors in* $\mathbb{R}^d$ *and* $\{r_i\}$*, a sequence of positive numbers. Suppose that there exists a* $c > 0$ *such that for all* i, j *we have* $\|v_i - v_j\| \geq c\sqrt{r_i r_j}$*. Then the set*

$$\left\{v \in \mathbb{R}^d \mid \exists \delta > 0 \text{ such that } \|v - v_i\| > \delta r_i \ \forall i\right\}$$

is an (α, β)*-winning set in* $\mathbb{R}^d$ *for all* $\alpha, \beta \in (0, 1)$ *such that* $1 - 2\alpha + \alpha\beta > 0$.

The case of badly approximable numbers in $\mathbb{R}$ falls out as a special case if we choose $\{x_i\}$ to be an enumeration of the rationals, and for $x_i = \frac{p}{q}$, where p and q are coprime integers and $q \neq 0$ ($q = 1$ if $p = 0$), choose $r_i = \frac{1}{q^2}$.

In place of the countable system of shrinking balls, whose complement is the set tested in the theorem for α-winning, analogous construction can be considered with countable systems of shrinking families of sets, including for example strips along affine subspaces of $\mathbb{R}^d$. An analogue of the theorem has been proved in this setting in [2]; in the general case the role of the radius, or rather diameter, is played by the thickness of the strip, or the set in general (for the latter "thickness" is defined by taking infimum over thicknesses of hyperplane strips containing the set). We shall not go into further details of the generalities here.

5.2 Toral Automorphisms

Let

$$\mathbb{T}^d \;=\; \left\{(z_1,\,\ldots,\,z_d)\;\mid\; z_j\in\mathbb{C},\;|z_j|=1\;\forall j=1,\ldots,d\right\}$$

be the d-dimensional torus. Any integral matrix $A=(m_{ij})$ with $\det A=\pm 1$ defines a continuous group automorphism T_A of $\mathbb{T}^d$, by

$$(z_1,\,\ldots,\,z_d)\;\longmapsto\;\left(z_1^{m_{11}}\,z_2^{m_{12}}\,\ldots\,z_d^{m_{1d}},\;\ldots,\;z_1^{m_{d1}}\,z_2^{m_{d2}}\,\ldots\,z_d^{m_{dd}}\right)$$

for all $(z_1,\,\ldots,\,z_d)\in\mathbb{T}^d$. Starting with any $z\in\mathbb{T}^d$ and applying an automorphism T_A repeatedly, we get the A-orbit $O_A(z)$ of z, namely $\{z,\;T_Az,\;T_A^2z,\;\ldots,\;T_A^kz,\;\ldots\}$. It is known that when no (complex) eigenvalue of A is a root of unity then the action of T_A as above is ergodic (see [11], for instance), and in particular it follows that for almost all $z\in\mathbb{T}^d$, $O_A(z)$ is dense in $\mathbb{T}^d$. There are of course points, such as those for which each coordinate is a root of unity, whose orbits are finite (and in particular not dense). However, the collection of all points with orbits that are not dense, which is a set of measure 0 for A satisfying the condition as above, defies simple description. The following may be noted in this context.

For any $d\times d$ matrix A, let $E(T_A)$ be the set of all $v=(v_1,\,\ldots,\,v_d)\in\mathbb{R}^d$ such that for $z=\left(e^{2\pi iv_1},\,\ldots,\,e^{2\pi iv_d}\right)$, $\overline{O_A(z)}$ does not contain any element of finite order; in particular $\overline{O_A(z)}$ is not dense in such a case.

Theorem 5.2 *For every A as above, $E(T_A)$ is an (α,β)-winning set for all $\alpha,\beta\in(0,1)$ such that $1-2\alpha+\alpha\beta>0$; thus it is α-winning for all $\alpha\in(0,\frac{1}{2}]$. Consequently, there exist uncountably many $z\in\mathbb{T}^d$ such that $\overline{O_A(x)}$ does not contain any element of finite order for any automorphism T_A of $\mathbb{T}^d$.*

The first part of the theorem was proved in [3] under an additional assumption that A is a semisimple matrix (viz., diagonalizable over the field of complex numbers); however the assumption turns out to be unnecessary for the proof, as has been clarified in [5]. As there are only countably many automorphisms T_A, the second part follows from the first, together with Theorem 3.2, and the fact that all α-winning subsets of $\mathbb{R}^d$ are uncountable.

5.3 Hyperbolic Geometry

Let M be a Riemannian manifold of dimension $d+1$ with constant negative curvature and finite Riemannian volume. Then M can be realized canonically as the quotient $\mathbb{H}^{d+1}/\Gamma$, where

$$\mathbb{H}^{d+1} \;=\; \left\{(x_1,\,\ldots,\,x_{d+1})\;\mid\; x_j\in\mathbb{R}\;\forall j=1,\ldots,d+1,\;\text{and}\;x_{d+1}>0\right\}$$

is the hyperbolic space equipped with the Riemannian metric $(dx_1^2 + \cdots + dx_{d+1}^2)/x_{d+1}^2$, and Γ is a group of isometries of $\mathbb{H}^{d+1}$.

We view $\mathbb{R}^d$, consisting of the subspace $\{x_1, \ldots, x_d, 0) \mid x_1, \ldots, x_d \in \mathbb{R}\}$ as the boundary of $\mathbb{H}^{d+1}$ in a natural way, when the latter is viewed as a subset of $\mathbb{R}^{d+1}$. The geodesics in $\mathbb{H}^{d+1}$ are semicircles with endpoints in $\mathbb{R}^d$, in the usual geometry of $\mathbb{H}^{d+1}$ as a subset of $\mathbb{R}^{d+1}$. In particular, every (positive time) geodesic trajectory $\{\gamma(t)\}_{t\geq 0}$ in $\mathbb{H}^{d+1}$ has a unique endpoint in $\mathbb{R}^d$. We note also that two such geodesic trajectories are asymptotic to each other if and only if the endpoints in $\mathbb{R}^d$ are the same.

The geodesic trajectories in $M = \mathbb{H}^{d+1}/\Gamma$ as above are just the images of geodesic trajectories in $\mathbb{H}^{d+1}$. A question of interest is to understand, when M is noncompact but has finite Riemannian volume, the class of geodesic trajectories in M which are bounded (viz. have compact closure in M). In this respect the following was deduced in [2] from theorem (5.1). When $d = 1$ and $\Gamma = \mathrm{SL}(2, \mathbb{Z})$, it corresponds to the result for badly approximable numbers, (Theorem 4.1).

Corollary 5.3 *Let M be a Riemannian manifold of dimension $d + 1$ with constant negative curvature and finite Riemannian volume. Let E be the set of $v \in \mathbb{R}^d$ such that v is an endpoint of a geodesic $\{\gamma(t)\}_{t\geq 0}$ in $\mathbb{H}^{d+1}$ whose image in M has compact closure in M. Then E is an α-winning set for $\alpha \in (0, \frac{1}{2}]$.*

6 Further Generalisations and Applications

The theory emerging from the developments described above has witnessed many generalizations and applications and it is beyond the scope of this article to discuss them in any detail. We shall content ourselves with some brief comments on the directions it has taken. Numerous authors have been involved in further generalizing the ideas indicated below, but we shall not go into all citations. The interested reader will be able to reach to the works through citations to some of the papers noted here.

6.1 Strong and Absolute Winning Sets

In [8], C.T. McMullen introduced two variations of the Schmidt game, on $\mathbb{R}^d$, $d \geq 1$, and the corresponding winning sets are known as *strong winning sets* and *absolute winning sets*. The former involves, like the Schmidt game two numbers $\alpha, \beta \in (0, 1)$ and the procedure for the game is analogous to the former, except that the conditions on the radii of $\{A_i\}$ and $\{B_i\}$ are now changed to $r(A_{i+1}) \geq \alpha r(B_i)$ and $r(B_i) \geq \beta r(A_i)$ for all $i \geq 0$. A winning set of the game is called (α, β)-*strong winning set* and a set which is (α, β)-strong winning for all $\beta \in (0, 1)$ is called α-*strong winning*.

In the other variation the players choose closed balls $\{A_i\}$ and $\{B_i\}$ such that for all $i \geq 0$, $B_{i+1} \subset B_i \backslash A_{i+1}$ and, for a fixed $\beta \in (0, \frac{1}{3})$, $r(B_{i+1}) \geq \beta r(B_i)$ and

$r(A_{i+1}) \leq \beta r(B_i)$. The game is called the absolute game and the winning sets are called *absolute winning sets*.

The strong winning sets and absolute winning sets share the properties of α-winning sets that we discussed in earlier sections, and moreover have the property that they are invariant under quasisymmetric homeomorphisms of $\mathbb{R}^d$, which is not true for general α-winning sets; a homeomorphism φ is said to be *k-quasisymmetric*, where k is a real number ≥ 1, if for any ball $B(x, r)$ in $\mathbb{R}^d$ there exists $s > 0$ such that $B(\varphi(x), s) \subset \varphi(B(x, r)) \subset B(\varphi(x), ks)$, and it is said to be *quasisymmetric* if it is k-quasisymmetric for some $k \geq 1$; it is known that when $d \geq 2$ the notions of quasisymmetric maps coincides with quasiconformal maps. Interesting applications of these ideas in hyperbolic geometry are found in [8].

6.2 Winning Sets on Lie Groups

D. Kleinbock and B. Weiss introduced a variation of the Schmidt games, on Lie groups; these do not involve a metric, as in the case of Schmidt games, but rather rely data arising group theoretically. It may be noted that in the case of $\mathbb{R}^d$ the collection of closed balls can be obtained by starting with a fixed ball B and applying to it all affine automorphisms of the form $v \longmapsto e^{-t}v + w$ for all $v \in \mathbb{R}^d$, where $t \in \mathbb{R}$ and $w \in \mathbb{R}^d$ are the parameters defining the affine automorphism; moreover, all balls smaller than B involve taking only positive t. Motivated by this, given a Lie group G the authors start with a fixed compact subset C of G with nonempty interior and consider the collection of all compact sets arising by affine automorphisms of G of the form $g \longmapsto \Phi_t(g)h$ for all $g \in G$, where $\{\Phi_t\}_{t\in\mathbb{R}}$ is a fixed one-parameter group of automorphisms of G which is contracting for positive t (viz., $\Phi_t(g) \to e$, the identity element, as $t \to \infty$), and $t \in \mathbb{R}$ and $h \in G$ define the family of transformations.

The role of the balls is now played by images of the set C under application of these affine automorphisms, with positive t. The procedure is then analogous, but involves some intricacies that we shall not go into. The details may be found in [7] (and the earlier papers of the authors cited there). Using the modified version of the game the authors proved in [7] a conjecture of G.A. Margulis on the abundance of certain kind of exceptional orbits of hyperbolic flows on homogeneous spaces.

6.3 Badly Approximable Numbers in Closed Subsets

Given that badly approximable numbers (and vectors in $\mathbb{R}^d$, $d \geq 2$) are abundant, in the light of the results on their winning nature for various games, one may wonder if we would be able to find them in say any perfect compact (or closed) subsets of $\mathbb{R}$. One immediate answer should be in the negative, since the complement of the set of badly approximable numbers is a set of positive Lebesgue measure and hence

contains perfect compact subsets, in fact of positive measure. It turns out however that under certain further conditions on the compact set such a conclusion is possible.

L. Fishman, D. Kleinbock and B. Weiss studied Schmidt games on fractals (see [6] and the references there), and it was shown that for supports of a class of measures, called "absolutely friendly measures" the intersection with the set of badly approximable numbers (respectively vectors in $\mathbb{R}^d$) is nonempty, and in fact has Hausdorff dimension equal to that of the support. An analogue of this for the case of automorphisms of tori discussed in (Sect. 7.5.2) was also proved in [1].

In [5] a variation of the Schmidt game was introduced on compact subsets of $\mathbb{R}^d$ satisfying certain regularity conditions; we shall not go into the technical details of the conditions involved, but mention only that totally disconnected subsets satisfying the conditions can be constructed, in abundance, emulating the construction of the classical Cantor set; as with the construction of the Cantor set, open intervals are removed from the pieces obtained at each stage, with controls on their sizes; in particular the conditions hold for all translates of the classical Cantor set in $\mathbb{R}$; similarly various sets satisfying the conditions can be explicitly constructed in higher dimensions. It was proved in particular that for compact subsets satisfying the conditions the intersection with the set of badly approximable vectors is uncountable. Analogous result is also proved in [5] for intersections of compact totally disconnected sets with sets of exceptional orbits of toral automorphisms discussed in (Sect. 7.5.2).

References

1. Broderick, R., Fishman, L., & Kleinbock, D. (2011). Schmidt's game, fractals, and orbits of toral endomorphisms. *Ergodic Theory and Dynamical Systems, 31*, 1117–1195.
2. Dani, S. G. (1986). Bounded orbits of flows on homogeneous spaces. *Commentarii Mathematici Helvetici, 61*, 636–660.
3. Dani, S. G. (1988). On orbits of endomorphisms of tori and the Schmidt game. *Ergodic Theory and Dynamical Systems, 8*, 523–529.
4. Dani, S. G. (1989). On badly approximable numbers, Schmidt games and bounded orbits of flows. In M. M. Dodson & J. A. G. Vickers (Eds.), *Number theory and dynamical systems*, London mathematical society lecture notes (Vol. 134). Cambridge: Cambridge University Press.
5. Dani, S. G., & Shah, H. (2012). Badly approximable numbers and vectors in Cantor-like sets. *Proceedings of the American Mathematical Society, 140*, 2575–2587.
6. Fishman, L. (2009). Schmidt's game on fractals. *Israel Journal of Mathematics, 171*, 77–92.
7. Kleinbock, D., & Weiss, B. (2013). Modified Schmidt games and a conjecture of Margulis. *Journal of Modern Dynamics, 7*, 429–460.
8. McMullen, C. T. (2010). Winning sets, quasiconformal maps and Diophantine approximation. *Geometric and Functional Analysis, 20*, 726–740.
9. Schmidt, W. M. (1966). On badly approximable numbers and certain games. *Transactions of the American Mathematical Society, 123*, 178–199.
10. Schmidt, W. M. (1980). *Diophantine approximation*. Berlin: Springer.
11. Walters, P. (1982). *An introduction to ergodic theory*, Graduate texts in mathematics. New York: Springer.