Ergodic Theory

Existence of Invariant and Ergodic measures.

W.A.Lakmi Niwanthi

Supervisor : Prof.S. Luzzatto

The Absus Salam International Center for Theoretical Physics
Strada Costiera 11, Miramare P.O.Box 586
I-34100 Trieste, Italy

Acknowlegement

I am grateful to my supervisor Prof. S. Luzzatto for his guidance, encouragement and patience during the course of the entire period of this Thesis.
I would also like to thank the Mathematics section for their contribution towards my academic progress and also the development of my research skills. Thanks also to the Director of ICTP, the partners, administrative body of ICTP and the entire community for providing a wonderful environment for learning throughout the diploma programme.
I wish to thank my parents and siblings for the undivided support, encouragement and inspiration for my career. My gratitude also goes to all my friends and everyone who contributed by supporting my work.

To my loving parents...

Contents

1 Preliminaries

1.1 Introduction

In this report, I will first give a brief description of ergodic theory which is related to dynamical system. Then explain in detail the problem on an existence of invariant, ergodic measure for continuous map and full branch maps.

Dynamical systems are used to model physical phenomena whose state change over time, for instance, in financial and economic forecasting, medical diagnosis, environment modeling etc.
In Ergodic theory, we describe the orbit of a point using measures. Our interest is to know whether we can describe the orbit of a point for a given function using measures. Birkhoff's Ergodic theorem plays an important role in such a study. It says that if there exists a measure which is invariant and ergodic for the given map, then we can describe the asymptotic distribution of a.e. points in the space w.r.t. that measure. Further more if that measure is absolutely continuous w.r.t. lebesgue measure then we are able to describe the asymptotic distribution of a large set of points with positive lebesgue measure.
In this project, I have studied existence of such a measure for continuous maps and full branch maps. Further more, I have studied that for a full branch map there can be uncountable measures which are invariant and ergodic. But there is only one measure which is absolutely continuous w.r.t. lebesgue measure.

Now let's talk about some concepts which will be useful in next chapters.

1.2 Topology

Definition 1.1. A subset K of a vector space X is said to be *convex* if, whenever it contains x and y, it also contains $\lambda x + (1-\lambda)y$ for $0 \leq \lambda \leq 1$. If K be a convex subset of a vector space X, a point x in K is called an *extreme point* if it is not an interior point of any line segment lying in K.i.e. x is extreme iff whenever $x = \lambda y + (1-\lambda)z$ with $0 < \lambda < 1$, $y \notin K$ or $z \notin K$. The intersection of all convex sets containing K is called *convex hull* of K.

For the proof of theorems 1.1, 1.2, 1.4 and 1.6 see [4].

Theorem 1.1 (Krein-Milman Theorem)**.** *Let K be a compact convex set in a locally convex topological vector space X. Then K is the closed convex hull of its extreme points.*

1.3 Measure Theory

Definition 1.2. A collection $\mathscr{A}$ of subsets of X is called an *algebra* of sets if whole set X is in $\mathscr{A}$, union of two sets of $\mathscr{A}$ is in $\mathscr{A}$ and complement of any set of $\mathscr{A}$ is again in $\mathscr{A}$. If any union of a countable collection of sets from $\mathscr{A}$ is again in $\mathscr{A}$, then algebra $\mathscr{A}$ is called σ- algebra. Then we called $(X, \mathscr{A})$ is a measurable space. The collection $\mathscr{B}(X)$ of *Borel sets* is the smallest σ- algebra which contains all of the open sets of X.

Definition 1.3. An *outer measure* μ^* on $(X, \mathscr{A})$ is a nonnegative extended real valued set function defined on all subsets of a space X with properties $\mu^*(\emptyset) = 0$, $A \subseteq B$ implies $\mu^*(A) \leq \mu^*(B)$ for $A, B \in \mathscr{A}$ and $E \subseteq \bigcup_{i=1}^{\infty} E_i$ implies $\mu^*(A) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$ for $\{E_i\}_{n\in\mathbb{N}}$.

Definition 1.4. We say set E *measureble* if for each set A, we have $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$. We denote measure of a measureble set by $\mu(A)$ and $\mu(A) = \mu^*(A)$. An extended real value function f is said to be *measureble* if its domain is measurable and for each measurable set B, $f^{-1}(B)$ is measurable.
Then $(X, \mathscr{A}, \mu)$ is called a measure space.

Definition 1.5. We say μ is a *probability measure* on X, if $\mu(X) = 1$ and $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ whenever $\{A_n\}_{n=1}^{\infty}$ is a sequence of members of $\mathscr{A}$ which are pairwise disjoint subsets of X. Let's denote all probability measures on X by $\mathcal{M}$.

Example 1.1.

The *dirac delta measure* is a probability measure defined for a given $p \in X$ and $A \in \mathscr{B}(X)$ by

$$\delta_p(A) = \begin{cases} 1 & \text{if } p \in A \\ 0 & \text{if } p \notin A \end{cases} \tag{1.1}$$

Definition 1.6. Let μ and υ are two measures defined on $(X, \mathscr{B}$. Then μ and υ are *mutually singular* if there is a non empty subset of X, E such that $\mu(E) = \mu(X)$ and $\upsilon(E) = 0$. The measure υ is said to be *absolutely continuous* with respect to measure μ (denote as $\upsilon \ll \mu$), if $\upsilon(A) = 0$ for each set $A \in \mathscr{B}$ for which $\mu(A) = 0$.
A property is said to hold *almost everywhere* (a.e.) if the set of points where it fails to hold is a set of measure zero.

Theorem 1.2 (Radon-Nikodym Theorem)**.** *Let $(X, \mathscr{B}, \mu)$ be a σ- finite measure space and let υ be a measure defined on $\mathscr{B}$ which is $\upsilon \ll \mu$. Then there is a nonnegative measurable function f such that for each set $E \in \mathscr{B}$ we have $\upsilon(E) = \int_E f d\mu$.*
The function f is unique in the sense that if g is any measurable function with this property then $g = f$ a.e.(μ).

The function f is called density of υ or Radon Nikodym derivative w.r.t υ. We denoted this density by $f = \dfrac{d\upsilon}{d\mu}$.
Now let's see some examples for mutually singular measures and absolutely continuous measures.

Example 1.2.

Define $\upsilon(E) = \int_E f d\mu$. Then for $f > 0$, $\upsilon \ll \mu$. Let $x \in X$. Then $\delta_x(\{x\}) = 1$ and $m(\{x\}) = 0$. Thus δ_x and lebesgue measure, m are mutually singular. But δ_x and lebesgue measure are not absolutely continuous w.r.t each other.

Definition 1.7. The weak star topology on $\mathcal{M}$ is the smallest topology making each of the maps $\mu \to \int f d\mu$ continuous for each $f \in C^\circ(M)$. In the weak star topology sequence of measures $\{\mu_n\}$ converges to measure μ in $\mathcal{M}$ iff for each $f \in C^\circ(M)$, $\int f d\mu_n$ converges to $\int f d\mu$.

See [3] for the proof of the following theorem.

Theorem 1.3. *Let m, μ be two Borel probability measures on the metric space X. If $\int f dm = \int f d\mu$ for each $f \in C^{\circ}(X)$, then $m = \mu$.*

1.4 Integration and Functional Analysis

Theorem 1.4 (Monotone Convergence Theorem)**.** *Suppose $\{f_n\}$ is an increasing sequence of integrable real valued functions on $(X, \mathscr{B}, \mu)$. If $\{\int f_n d\mu\}$ is a bounded sequence of real numbers then $\lim_{n\to\infty} f_n$ exists a.e. and is integrable and $\displaystyle\int \lim f_n d\mu = \lim \int f_n d\mu$.*

Definition 1.8. A family $\mathscr{F}$ of continuous functions from a metric space (X, d) to a metric space (Y, σ) is called *equicontinuous* at the point $x \in X$ if for given $\epsilon > 0$ there is an open set O containing x such that $\sigma(f(x), f(y)) < \epsilon$ for all $y \in O$ and all $f \in \mathscr{F}$. The family $\mathscr{F}$ is said to be equicontinuous if it is equicontinuous at each point $x \in X$.

Refer [3] for the proof of Ascoli-Arzela Theorem.

Theorem 1.5 (Ascoli-Arzela Theorem)**.** *Let $\mathscr{F}$ be an equicontinuous family of functions on a seperable space X. Then each sequence $\{f_n\}$ in $\mathscr{F}$ which is bounded at each point (of a dense subset) has a subsequence $\{f_{n_k}\}$ that converges pointwise to a continuous function, the convergence being uniform on each compact subset of X.*

Theorem 1.6 (Riesz Representation Theorem)**.** *Let M be a compact Hausdorff space and $C^{\circ}(M)$ the space of continuous real valued functions on M. Then to each bounded linear functional F on $C^{\circ}(M)$ there corresponds a unique finite signed Baire measure μ on M such that $F(f) = \int f d\mu$, for each f in $C(M)$.*

Notice:
Let $\mu \in \mathcal{M}$. Define the functional $\varphi_\mu : C^0(M, \mathbb{R}) \to \mathbb{R}$ by $\varphi_\mu(f) = \int f d\mu$. Then, $|\varphi_\mu(f)| = |\int f d\mu| \leq \int |f| d\mu = L < \infty$ since f continuous and M compact. Therefore φ_μ is bounded. Also φ_μ is linear. For let $\lambda_1, \lambda_2 \in \mathbb{R}$ and $f_1, f_2 \in C^0(M, \mathbb{R})$. Then $\varphi_\mu(\lambda_1 f_1 + \lambda_2 f_2) = \int (\lambda_1 f_1 + \lambda_2 f_2) d\mu = \lambda_1 \int f_1 d\mu + \lambda_2 \int (f_2) d\mu = \lambda_1 \varphi(f_1) + \lambda_2 \varphi(f_2)$. And for $f \in C^0(M, \mathbb{R}^+)$, $\varphi_\mu(f) \geq 0$. Also $\varphi_\mu(I) = \int I(x) = \mu(M) = 1$, where $I(x) = 1$ for each $x \in M$.

Now we proceed to Ergodic theory which has been the main focus of this project.

1.5 Dynamical system

In mathematical point of view, the theory of dynamical system is the study of the global orbit structure of maps and flows. Conjugacies help to describe the orbit of a map using orbit of an another map.

Definition 1.9. Let X, Y be two topological spaces and $f : X \to X$, $g : Y \to Y$ are continuous maps. We say that f and g are *topologically conjugate* if there exists a homeomorphism $h : X \to Y$ such that $h \circ f = g \circ f$.

Conjugacy maps orbits of f to orbits of g. i.e. $f^n(x)$ map to $g^n(h(x))$ for each $x \in X$ and $n \in \mathbb{N}$. This is an equivalence relation on the space of continuous maps defined on a topological space. Therefore each equivalence class contains all functions which has the same dynamics in topological view point.

For example let's define a topological conjugate between two maps, f_p and f as below. We shall use this in chapter 3, to show that existence of uncountable ergodic, invariant measures for f. We sketch the proof that all these maps are topologically conjugate to each other.

Example 1.3.

Let $p \in (0,1)$. Define the map $f_p : [0,1] \to [0,1]$ by,

$$f_p(x) = \begin{cases} \frac{x}{p} & \text{if } x \in [0,p) \\ \frac{x}{1-p} - \frac{p}{1-p} & \text{if } x \in [p,1] \end{cases} \tag{1.2}$$

By using symbolic coordinates we can represent each $x \in [0,1)$ uniquely such that $x = x_0x_1x_2\cdots$ where,

$$x_i = \begin{cases} 0 & \text{if } f_p^i(x) \in [0,p) \\ 1 & \text{if } f_p^i(x) \in (p,1] \end{cases} \tag{1.3}$$

And let $f(x) = 2x(\mod 1)$ on $[0,1]$. This is the case when $p = \frac{1}{2}$. Therefore in the same way we can give symbolic coordinates to each $x \in [0,1]$ using map f. Then by mapping each points which has same symbolic coordinates and preimages of p to preimages of $\frac{1}{2}$ we can define a homeomorphism, say h_p on $[0,1]$. Therefore f_p and f are topologically conjugate by h_p.

To describe an orbit, we define the basin of a measure μ for the map f by,

$$\mathscr{B}_\mu = \left\{ x \in M : \frac{1}{n}\sum_{i=0}^{n-1} \varphi \circ f^i(x) \to \int \varphi d\mu \text{ for each } \varphi \in C^0(M,\mathbb{R}) \right\}.$$

Then we look for measures such that $\mathscr{B}_\mu$ contains many x from M, i.e. $m(\mathscr{B}_\mu) > 0$. Because asymptotic distribution of any $x \in \mathscr{B}_\mu$ can described by the measure μ. However every function does not have a basin with positive lebesgue measure, the identity map is an example. Birkhoff's Ergodic theorem gives a method to find such a measure in the way that if a measure is invariant, ergodic then $\mu(\mathscr{B}_\mu) = 1$. Also if measure μ is absolutely continuous with respect to lebesgue measure then $m(\mathscr{B}_\mu) > 0$. Therefore the asymptotic distribution of all of the many points in the space are able to describe.

In next two sections, we present two facts; invariant and ergodicity that will be the key definitions for next three chapters.

1.6 Measure Preserving Transformations

Definition 1.10. Suppose $(X_1, \mathscr{B}_1, \mu_1), (X_2, \mathscr{B}_2, \mu_2)$ are probability spaces. A measurable transformation $T : X_1 \to X_2$ is *measure preserving (invariant)* if $\mu_1(T^{-1}(B_2) = \mu_2(B_2)$ for each $B_2 \in \mathscr{B}_2$. Let's denote all invariant measures for map T on $(X, \mathscr{B})$ by $\mathcal{M}_{\mathcal{T}}$. We shall show that $\mathcal{M}_{\mathcal{T}}$ is convex and compact in weak star for continuous maps defined on a compact space. For example any probability measure is T-invariant for identity map T.

1.7 Ergodic Theory

Ergodicity describe a dynamical system which has same behaviour averaged over time as averaged over the space of the system states.

Definition 1.11. Let $(X, \mathscr{B}, \mu)$ be a probability measure space. A measurable transformation T is called *ergodic* if $B \in \mathscr{B}$ with $T^{-1}(B) = B$ satisfy $\mu(B) = 0$ or $\mu(B) = 1$.

Example 1.4.

Let $T : X \to X$ be a measurable function. Then the dirac delta, δ_p is T invariant and ergodic where p is a fixed point for T.

Example 1.5.

Let I be an interval of $\mathbb{R}$. If $T : I \to I$ be a piecewice affine full branch map. Then lebesgue measure, m is invariant and ergodic for T.

We use the following two theorems to explain some concepts and examples in next chapters.

Theorem 1.7 (Poincare Recurrence Theorem). *Let $T : X \to X$ be a measure preserving transformation of a probability space $(X, \mathscr{B}, \mu)$. Let $E \in \mathscr{B}$ with $\mu(E) > 0$. Then almost all points of E return infinitely often to E under positive iteration by T.*

The proof of Poincare Recurrence theorem, can find in [3].

Theorem 1.8 (Birkhoff's Ergodic Theorem). *Let $(M, \mathbb{B}, \mu)$ be a probability space, $T : M \to M$ be a measure preserving transformation and $f \in L^1(M, \mathscr{B}, \mu)$. Then, $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x)$ exists a.e.*

If $A \in \mathscr{B}$ with $T^{-1}(A) = A$ (i.e. if μ is ergodic), then $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) = \int_A f d\mu$ a.e.

See [2] for the proof of the Birkhoff's Ergodic Theorem.

It's not immediate that, a measure is ergodic for a given map T. But the following lemma says that any measure is ergodic for T if it is absolutely continuous w.r.t an another measure which is ergodic for map T.

Lemma 1.1. *Let $f : I \to I$ be a measurable map and let μ_1 and μ_2 be two probability measures with $\mu_1 \ll \mu_2$ and μ_2 is ergodic for f. Then μ_1 is ergodic for f.*

Proof. Let $A \in \mathscr{B}(I)$ such that $\mu_1(A) > 0$. Since $\mu_1 \ll \mu_2$, $\mu_2(A) > 0$. Then $\mu_2(A) = 1$ as μ_2 is ergodic for f. Therefore $\mu_2(I\ A) = 0$ and hence $\mu_1(I\ A) = 0$ by absolutely continuity of μ_1 w.r.t. μ_2. Thus $\mu_1(A) = 1$. i.e. μ_1 is ergodic for f.

□

Definition 1.12. Let $I \subset \mathbb{R}$ be an interval. A map $f : I \to I$ is *a full branch map* if there exists a finite or countable partition $\mathscr{P}$ of I, into subintervals such that for each $w \in \mathscr{P}$ the map $f|_{int(w)} : int(w) \to int(I)$ is a bijection.

If for each $w \in \mathscr{P}$ the map $f|_{int(w)} : int(w) \to int(I)$ is a homeomorphism (resp. C^1 diffeomorphism, C^2 diffeomorphism, affine) then f is called *a piecewise continuous* (resp. C^1, C^2, affine) full branch map.

A full branch map has *bounded distortion* if $\sup_{n\geq 1} \sup_{w^{(n)}\in\mathscr{P}^{(n)}} \sup_{x,y\in w^{(n)}} \log\left|\frac{Df^n(x)}{Df^n(y)}\right| < \infty$.

Theorem 1.9. *Let $f : I \to I$ be a full branch map with* $\sup_{n\geq 1} \sup_{w^{(n)}\in\mathscr{P}^{(n)}} \sup_{x,y\in w^{(n)}} \log\left|\frac{Df^n(x)}{Df^n(y)}\right| < \infty$. *Then the Lebesgue measure, m is ergodic.*

2 Existence of Invariant and Ergodic measure for continuous functions on a compact space

Before we show the existence of invariant and ergodic measure for a continuous functions on a compact space, we define a measure called push forward, which will be used here and in next two chapters.

2.1 Push forward measure

Let T be a invertible map from space M to itself. Then define $T_* : \mathcal{M} \to \mathcal{M}$, by $T_*\mu(A) = \mu\left(T^{-1}(A)\right)$ for $\mu \in \mathcal{M}$ and $A \in \mathscr{B}(M)$. Then $T_*\mu$ is called the push forward of μ by T. Then $T_*\mu$ is a probability measure since μ is a probability measure on M. Also μ is invariant iff $T_*\mu = \mu$.

Let $\upsilon \in \mathcal{M}$ and define $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} T_*^k \upsilon$.

Then μ_n is a probability measure, since $\mu_n(X) = \frac{1}{n}\sum_{k=0}^{n-1} T_*^k \upsilon(X) = \frac{1}{n}\sum_{k=0}^{n-1} 1 = 1$ and for $\{A_i\}_{i\in\mathbb{N}}$ disjoint measurable subsets of X,

$$\begin{aligned}
\mu_n\left(\bigcup_{i\in\mathbb{N}} A_i\right) &= \frac{1}{n}\sum_{k=0}^{n-1} T_*^k \upsilon\left(\bigcup_{i\in\mathbb{N}} A_i\right) \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\left(\sum_{i\in\mathbb{N}} T_*^k \upsilon(A_i)\right) \\
&= \sum_{i\in\mathbb{N}}\left(\frac{1}{n}\sum_{k=0}^{n-1} T_*^k \upsilon(A_i)\right) \\
&= \sum_{i\in\mathbb{N}} \mu_n(A_i)
\end{aligned}$$

The following examples will help to understand the convergence of $\{\mu_n\}_{n\in\mathbb{N}}$.

Example 2.1.

Let $f(x) = x^2$ on $[0,1]$. We shall show that μ_n converges in weak star to measure δ_0. Let $a \in [0,1]$ and $A = [0,a]$. Now consider the measure $\overline{\mu_n}(A) = f_*\upsilon(A) = \upsilon\left(f^{-n}(A)\right)$ for $n \in \mathbb{N}$.

Then,

$$\begin{aligned}
\overline{\mu_1}(A) &= f_*\upsilon(A) = \upsilon\left(f^{-1}([0,a])\right) = \upsilon\left(\left[0, a^{\frac{1}{2}}\right]\right) \\
\overline{\mu_2}(A) &= f_*\upsilon(A) = \upsilon\left(f^{-2}([0,a])\right) = \upsilon\left(\left[0, a^{\frac{1}{2^2}}\right]\right) \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
\overline{\mu_n}(A) &= f_*\upsilon(A) = \upsilon\left(f^{-2}([0,a])\right) = \upsilon\left(\left[0, a^{\frac{1}{2^n}}\right]\right).
\end{aligned}$$

Then $\lim_{n\to\infty}\overline{\mu_n}(A) = \lim_{n\to\infty}\upsilon([0,a^{\frac{1}{2^n}}]) = \upsilon([0,1]) = 1$. Since a is arbitrary, for any measurable set B on $[0,1]$ which contains 0, $\lim_{n\to\infty}\overline{\mu_n}(B) = 1$. For that $\bigcap_{n\in\mathbb{N}}\left[0,\frac{1}{2^n}\right] \subseteq B$ and $\lim_{n\to\infty}\overline{\mu_n}\left(\left[0,\frac{1}{2^n}\right]\right) = \overline{\mu_n}\left(\bigcap_{n\in\mathbb{N}}\left[0,\frac{1}{2^n}\right]\right) \leq \overline{\mu_n}(B)$. Since for each $n\in\mathbb{N}$, $\overline{\mu_n}\left(\left[0,\frac{1}{2^n}\right]\right) = 1$, $\overline{\mu_n}(B) = 1$. Therefore $\overline{\mu_n}$ converges to δ_0. Then the average of $\overline{\mu_n}$, μ_n converges to δ_0.

Remark: δ_0 is an invariant, ergodic measure for f, and is not absolutely continuous w.r.t m. But $\mathscr{B}_{\delta_0} > 0$ since $\frac{1}{n}\sum_{i=0}^{n-1}\varphi\circ f(x)$ converges to $\int\varphi d\mu$ for each $x\in[0,1)$ and for $\varphi\in C^\circ([o,1])$.

Example 2.2.

Consider the map $f(x) = 10x(\mod 1)$ on $[0,1]$. Let $\underline{p}$ =.ppp$\cdots$, $\underline{q}$ =.qqq$\cdots$ and $\underline{a}$ =.pqppqqpppqqq$\cdots$ where, number of consequence p's and q's are increase. Then $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1}f_*^n\delta_{\underline{a}} = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{\underline{a}}\circ f^{-i} = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{f^i(\underline{a})}$ converges to $\frac{1}{2}\left(\delta_{\underline{p}}+\delta_{\underline{q}}\right)$. To see this let $\epsilon > 0$. Then there exist $n_1, n_{p_1}\in\mathbb{N}$ such that $f^{n_1+k}(\underline{a})\in B_\epsilon(\underline{p})$ for $k = 1,2,3,\cdots n_{p_1}$. [1] Similarly there exist $n_2, n_{q_1}\in\mathbb{N}$ such that $n_2 \geq n_1 + n_{p_1}$ and $f^{n_2+k}(\underline{a})\in B_\epsilon(\underline{q})$ for $k = 1,2,3,\cdots n_{q_1}$.
By continuing this we can find two sequences $\{n_{p_i}\}\subseteq\mathbb{N}$ and $\{n_{q_i}\}\subseteq\mathbb{N}$ such that $f^i(\underline{a})$ stay in $B_\epsilon(\underline{p})$, n_{p_1} number of time and then n_{q_1} number of time in $B_\epsilon(\underline{q})$. Again n_{p_2} number of time in $B_\epsilon(\underline{p})$ and then n_{q_2} number of time in $B_\epsilon(\underline{q})$. Since number of consequence p's and q's are increase in $\underline{a}$, sequences $\{n_{p_i}\}$ and $\{n_{q_i}\}$ are increasing sequences. Therefore $f^i(\underline{a})$ stay mostly in $B_\epsilon(\underline{p})$ and $B_\epsilon(\underline{q})$ alternately as i increases.

Then for any $A\in\mathscr{B}([0,1])$ with $B_\epsilon(\underline{p})\bigcap B_\epsilon(\underline{q})\subseteq A$, $\mu_n(A) = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{f^i(\underline{a})}(A)$ converges to 1 as when we iterate $\underline{a}$, for many $i\in\mathbb{N}$, $f^i(\underline{a})$ remains in $B_\epsilon(\underline{p})$ and $B_\epsilon(\underline{q})$. If only $B_\epsilon(\underline{p})\subseteq A$, $\mu_n(A)$ converges to $\frac{1}{2}$. Therefore $\mu_n(A) = \frac{1}{n}\sum_{i-0}^{n-1}f_*^i\delta_{\underline{a}}(A)$ converges to $\frac{1}{2}(\delta_{\underline{p}}+\delta_{\underline{q}})$.

Lemma 2.1. *For all $f\in L^1(\upsilon)$, $\int f d(T_*\upsilon) = \int f\circ T d\upsilon$.*

Proof. Let $A\in\mathcal{B}(\mathcal{M})$. Consider characteristic function χ_A. Then $\int\chi_A d(T_*\upsilon) = T_*\upsilon(A) = \upsilon\left(T^{-1}(A)\right) = \int\chi_{T^{-1}(A)}d\upsilon = \int\chi_A\circ T d\upsilon$. Thus the result is true for simple funtions on M. Now suppose f is a measurable positive function on M. Then there exists monotonically increasing, sequence of simple functions $\{f_n\}$ on M such that $\lim_{n\to\infty}f_n = f$. Then

[1]For more explanation suppose $\epsilon = .01$, $\underline{p} = .222\cdots$, $\underline{q} = .333\cdots$ and $\underline{a} = .232233222\cdots$. Then $\left|\underline{p} - f^2(\underline{a})\right| = |.222\cdots - .2233222\cdots| = .0011\cdots < \epsilon$. Therefore $f^2(\underline{a})\in B_\epsilon(\underline{p})$. But $f^3(\underline{a})\notin B_\epsilon(\underline{p})$ as $\left|\underline{p} - f^3(\underline{a})\right| = .011\cdots > \epsilon$. Therefore $n_1 = 2$ and $n_{p_1} = 1$.
Again $\left|\underline{p} - f^4(\underline{a})\right| = |.333\cdots - .33222333\cdots| = .0011\cdots < \epsilon$. Thus $f^4(\underline{a})\in B_\epsilon(\underline{q})$. Since $\left|\underline{q} - f^5(\underline{a})\right| = .011\cdots > \epsilon$, $f^3(\underline{a})\notin B_\epsilon(\underline{p})$. Therefore $n_2 = 4$ and $n_{q_1} = 1$.
Similarly $\left|\underline{p} - f^6(\underline{a})\right|$ and $\left|\underline{p} - f^7(\underline{a})\right|$ less than ϵ, but $\left|\underline{p} - f^8(\underline{a})\right| > \epsilon$. Thus $n_3 = 6$ and $n_{p_2} = 2$ and etc.

$\int f_n d\left(T_* \upsilon\right) = \int f_n \circ T d\upsilon$ for each $n \in \mathbb{N}$. Therefore $\lim_{n\to\infty} \int f_n d\left(T_* \upsilon\right) = \lim_{n\to\infty} \int f_n \circ T d\upsilon$.
Therefore by Monotone Convergence theorem, $\int f d\left(T_* \upsilon\right) = \int f \circ T d\upsilon$ for positive measurable functions defined on M. Since any function f, can write as $f = f^+ - f^{-1}$ with $f^+, f^{-1} \geq 0$, result is hold for any measurable function f on M.
Therefore for all $f \in L^1(\upsilon)$, $\int f d\left(T_* \upsilon\right) = \int f \circ T d\upsilon$. □

Lemma 2.2. *υ is invariant iff for any $f \in C^\circ(M, \mathbb{R})$, $\int f \circ T d\upsilon = \int f d\upsilon$.*

Proof. $\Rightarrow$

Suppose υ is invariant. Let $f \in C^\circ(M, \mathbb{R})$ and $A \in \mathcal{B}(M)$.
Then $T_*\mu(A) = \mu\left(T^{-1}(A)\right) = \mu(A)$. Therefore $\int f d\mu = \int f dT_*\mu = \int f \circ T d\mu$, by lemma 2.1. Hence $\int f \circ T d\mu = \int f d\mu$ for each $f \in C^\circ(M, \mathbb{R})$.

$\Leftarrow$

Now suppose for any $f \in C^\circ(M, \mathbb{R})$, $\int f \circ T d\upsilon = \int f d\upsilon$.
Then $\int f d(T_*\mu) = \int f \circ T d\mu$, by lemma 2.1. Therefore $\int f dT_*\upsilon = \int f d\upsilon$. i.e. $\varphi_{T_*\mu}(f) = \varphi_\mu(f)$. Then by Riesz Representation theorem, $T_*\mu = \mu$. Hence μ is invariant. □

We state the following proposition which we shall use in the proof of next lemma. See [5] for the proof.

Proposition 2.1. *$\mathcal{M}$ is weak star compact if space M is compact.*

Lemma 2.3. *$\mathcal{M}_\mathcal{T}$ is convex and weak star compact.*

Proof. Let $\mu_1, \mu_2 \in \mathcal{M}_\mathcal{T}$ and $t \in [0,1]$. Let $A \in \mathcal{B}(M)$. Consider $\mu(A) = t\mu_1(A) + (1-t)\mu_2(A)$. Then $\mu(X) = t\mu_1(X) + (1-t)\mu_2(X) = t + (1-t) = 1$. And let $\{A_k\}_{k\in\mathbb{N}} \subset \mathbb{B}$. Then

$$\begin{aligned}\mu\left(\bigcup_{k\in\mathbb{N}} A_k\right) &= t\mu_1\left(\bigcup_{k\in\mathbb{N}} A_k\right) + (1-t)\mu_2\left(\bigcup_{k\in\mathbb{N}} A_k\right) \\ &= t\sum_{k\in\mathbb{N}} \left(\mu_1(A_k)\right) + (1-t)\sum_{k\in\mathbb{N}} \left(\mu_2(A_k)\right) \\ &= \sum_{k\in\mathbb{N}} \left(t\mu_1(A_k) + (1-t)\mu_2(A_k)\right) = \sum_{k\in\mathbb{N}} \mu(A_k)\end{aligned}.$$

Also $\mu\left(T^{-1}(A)\right) = t\mu_1\left(T^{-1}(A)\right) + (1-t)\mu_2\left(T^{-1}(A)\right) = t\mu_1(A) + (1-t)\mu_2(A) = \mu(A)$.
Hence $\mathcal{M}_\mathcal{T}$ is convex.

To show $\mathcal{M}_\mathcal{T}$ is weak star closed subset of $\mathcal{M}$. For let $\{\mu_n\} \subseteq \mathcal{M}_\mathcal{T}$ such that μ_n converges in weak stat to $\mu \in \mathcal{M}$. Let $f \in C^0(M, \mathbb{R})$. Then $\int f \circ T d\mu = \lim_{n\to\infty} \int f \circ T d\mu_n = \lim_{n\to\infty} \int f d\mu_n = \int f d\mu$.

Therefore μ is T invariant. Thus $\mathcal{M}_\mathcal{T}$ is closed in $\mathcal{M}$.
Therefore since $\mathcal{M}$ is weak star compact, $\mathcal{M}_\mathcal{T}$ weak star compact.

□

2.2 Existence of invariant measure for continuous maps on a compact space

Theorem 2.1. *If $T : M \to M$ is a continuous transformation of a compact metric space M then $\mathcal{M}_{\mathcal{T}}$ is non-empty.*

Proof. Now suppose space M is compact and T is a continuous map such that, $T : M \to M$. Then by proposition 2.1, probability space, $\mathcal{M}$ is compact in weak star topology. Therefore there exists a convergent subsequence $\{\mu_{n_k}\}_{n_k \in \mathbb{N}}$ for sequence $\{\mu_n\}_{n \in \mathbb{N}}$, where $\mu_n = \dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1} T_*^k \upsilon$. Say $\mu = \displaystyle\lim_{n_k \to \infty} \mu_{n_k}$. Let $f \in C^0(X, \mathbb{R})$.

Then, $\displaystyle\int f \circ T d\mu_{n_k} - \int f d\mu_{n_k} = \int (f \circ T - f) d\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} T_*^j \upsilon\right)$

By lemma 2.1 $\displaystyle\int f dT_*^j \upsilon = \int f \circ T^j d\upsilon$ and $\int f \circ T dT_*^j \upsilon = \int f \circ T^{i+1} d\upsilon$. Therefore,

$$\begin{aligned}\int (f \circ T - f) d\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} T_*^j \upsilon\right) &= \frac{1}{n_k}\int \sum_{j=0}^{n_k-1} \left(f \circ T^{j+1} - f \circ T^j\right) d\upsilon \\ &= \int (f \circ T^{n_k} - f) d\upsilon\end{aligned}$$

Then,

$$\begin{aligned}\lim_{n_k \to \infty}\left|\int f \circ T d\mu_{n_k} - \int f d\mu_{n_k}\right| &= \frac{1}{n_k}\left|\int (f \circ T^{n_k} - f)\right| \\ &\leq \frac{1}{n_k}\left(\int |f \circ T^{n_k}| d\upsilon + \int |f| d\upsilon\right)\end{aligned}$$

Since f is integrable, $\displaystyle\int |f| d\upsilon$ is finite.

Hence $\displaystyle\lim_{n_k \to \infty} \frac{1}{n_k}\left(\int |f \circ T^{n_k}| d\upsilon + \int |f| d\upsilon\right) = 0$.

Therefore $\displaystyle\lim_{n_k \to \infty}\left(\int f \circ T d\mu_{n_k} - \int f d\mu_{n_k}\right) = 0$. Thus $\displaystyle\int f \circ T d\mu = \int f d\mu$.

Hence $\mu \in \mathcal{M}_{\mathcal{T}}$, by lemma -2.2. □

Our next example shows that theorem does not hold if space M is not compact.

Example 2.3.

Let $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x + 1$. Assume f has an invariant probability measure μ. Let $A \in \mathscr{B}$ with $\mu(A) > 0$ and $a = \sup A$. Notice that we can clearly assume that A is bounded.

Let $x \in A$. Since f is an expansion on $\mathbb{R}$, there exists $n_0 \in \mathbb{N}$ such that $f^i(x) > a$ for each $i > n_0$. This contradicts the Poincare Recurrence theorem.

Therefore there are no invariant probability measure for f on $\mathbb{R}$.

This theorem also fails for discontinuous functions. See the following example.

Example 2.4.

First let's consider the map $f(x) = x/2$ on $[0, 1]$. Then δ_0 is the unique invariant measure for f. For that let $A \in \mathscr{B}([0, 1])$. If $0 \in A$ then $0 \in f^{-1}(A)$ and if $0 \notin A$ then $0 \notin f^{-1}(A)$. Therefore $\delta_0(A) = \delta_0(f^{-1})(A)$. Therefore δ_0 is f invariant. For uniqueness let's assume there exists an f invariant measure μ which is distinct from δ_0. Then there exists $A \in \mathscr{B}([0, 1])$ such that $0 \notin A$ and $\mu(A) > 0$. Let $\epsilon > 0$. Since $f^i(x)$ converges to 0, there exists $n_0 \in \mathbb{N}$ such that $f^i(x) < \epsilon$, for each $i \geq n_0$. i.e. for any $\epsilon > 0$ such that $A \subset (\epsilon, 1]$, $f^i(x) \notin A$ for each $i \geq n_0$. This contradicts the Poincare Recurrence theorem. Hence δ_0 is the unique invariant measure for f.

Consider the function g on $[0, 1]$ given by,

$$g(x) = \begin{cases} 1 & \text{if } x = 0 \\ x/2 & x \in (0, 1] \end{cases} \tag{2.1}$$

Then g is not continuous at 0. Assume δ_0 is g invariant. And let $A = \{0\}$. Then $\delta_0(A) > 0$ and $g^i(0) = \dfrac{1}{2^{i-1}}$ for each $i \in \mathbb{N}$. Then by Poincare Recurrence theorem there exists $\{n_k\}_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that $g^{n_k}(0) \in A$. i.e. $\dfrac{1}{2^{n_k}} = 0$ for each $n_k \in \mathbb{N}$. This is a contradiction. Hence δ_0 is not invariant for g. Therefore there is no invariant measure for g.

2.3 Existence of ergodic, invariant measure

We are interested in both invariant and ergodic measures. In this section, we prove that there exists atleast one ergodic, invariant measure for continuous maps defined on compact spaces.

Proposition 2.2. *$\mu \in \mathcal{M}_{\mathcal{T}}$ is ergodic if μ is an extremal point of $\mathcal{M}_{\mathcal{T}}$.*

Proof. Let μ be a T be an extremal point of $\mathcal{M}_{\mathcal{T}}$. Let $E \in \mathscr{B}(M)$ such that $T^{-1}(E) = E$. Assume that μ is not ergodic. Then $\mu(E) \neq 0$ and $\mu(E) \neq 1$.
Define $\mu_1 = \dfrac{\mu(E \bigcap A)}{\mu(E)}$ and $\mu_2 = \dfrac{\mu(E \bigcap A^c)}{\mu(E^c)}$ for $A \in \mathscr{B}(M)$.
Then μ_1 and μ_2 are probability measures as $\mu_1(X) = \dfrac{\mu(E \cap X)}{\mu(E)} = 1$ and for $\{A_k\}_{k\in\mathbb{N}}$ be disjoint measurable subsets of $\mathscr{B}(M)$,

$$\begin{aligned} \mu_1\left(\bigcup_{k\in\mathbb{N}} A_k\right) &= \frac{\mu\left(E \cap \left(\bigcup_{k\in\mathbb{N}} A_k\right)\right)}{\mu(E)} \\ &= \frac{\mu\left(\bigcup_{k\in\mathbb{N}} (A_k \cap E)\right)}{\mu(E)} \\ &= \sum_{k\in\mathbb{N}} \left(\frac{\mu(E \cap A_k)}{\mu(E)}\right) \\ &= \sum_{k\in\mathbb{N}} \mu_1(A_k) \end{aligned}.$$

Also μ_1 and μ_2 are invariant measures. For that let $A \in \mathscr{B}(M)$.
Then $\mu_1\left(T^{-1}(A)\right) = \dfrac{\mu\left(E \cap T^{-1}(A)\right)}{\mu(E)}$. Since we take E such that $T^{-1}(E) = E$,

$$\frac{\mu\left(E \cap T^{-1}(A)\right)}{\mu(E)} = \frac{\mu\left(T^{-1}(E) \cap T^{-1}(A)\right)}{\mu(E)}.$$

Also $\mu\left(T^{-1}(E \cap A)\right) = \mu(E \cap A)$ since μ is an invariant measure.

Therefore,

$$\frac{\mu\left(E \cap T^{-1}(A)\right)}{\mu(E)} = \frac{\mu\left(T^{-1}(E) \cap T^{-1}(A)\right)}{\mu(E)} = \frac{\mu\left(T^{-1}(E \cap A)\right)}{\mu(E)} = \frac{\mu(E \cap A)}{\mu(E)} = \mu_1(A).$$

Thus $\mu_1\left(T^{-1}(A)\right) = \mu_1(A)$. Thus μ_1 is an invariant measure. In similarly μ_2 also an invariant measure.

Since $\mu_1(E) = 1$ and $\mu_2(E) = 0$ they are mutually singular. Therefore μ_1 and μ_2 are distinct.
Also $\mu(A) = \mu(E)\mu_1(A) + (1 - \mu(E))\mu_2(A)$ and $0 < \mu(E) < 1$. This contradicts that μ is an extremal point of $\mathcal{M}_\mathcal{T}$. Hence μ is ergodic.

□

Theorem 2.2. *There exists at least one ergodic, invariant measure for continuous functions on a compact space.*

Proof. By lemma 2.3 $\mathcal{M}_\mathcal{T}$ is a convex and weak star compact subset of $\mathcal{M}$. Then by Krein - Milmal theorem, $\mathcal{M}_\mathcal{T}$ is closed convex hull of it's extremal points. Since by Theorem 2.1 $\mathcal{M}_\mathcal{T}$ is non empty, there exist atleast one extremal point in $\mathcal{M}_\mathcal{T}$. This proves the existence of ergodic, T-invariant measure. □

3 Existence of uncountable number of measures

We can use measurable conjugacy of two maps, to look for an invariant and ergodic measure for a map. In this chapter, we use it to show that there can be uncountable number of invariant, ergodic measures for a function.

Definition 3.1. Let X, Y be two metric spaces and $f : X \to X$ and $g : Y \to Y$ be two maps. We say that f and g are *measurably conjugate* if there exists a bijection $h : X \to Y$ such that $h \circ f = g \circ h$ and h and h^{-1} are measurable.

Theorem 3.1. *Let X, Y be two metric spaces and $f : X \to X$ and $g : Y \to Y$ be two maps such that measurably conjugate by $h : X \to Y$. Then,*

1. *μ is invariant under f iff $h_*\mu$ is invariant under g*
2. *μ is ergodic under f iff $h_*\mu$ is ergodic under g*

Proof. 1. Suppose μ is invariant under f. Let $A \in \mathscr{B}(Y)$.
Then $h_*\mu(g^{-1}(A)) = \mu(h^{-1} \circ g^{-1}(A)) = \mu(h^{-1}(A)) = h_*\mu(A)$. Thus $h_*\mu$ is invariant under g.
Now suppose $h_*\mu$ is invariant under g and let $A \in \mathscr{B}(X)$. Then,

$$\begin{aligned}
\mu(f^{-1}(A)) &= \mu(h^{-1} \circ g^{-1} \circ h(A)) \\
&= h_*\mu(g^{-1} \circ h(A)) \\
&= h_*\mu(h(A)) \\
&= \mu(h^{-1} \circ h(A)) \\
&= \mu(A)
\end{aligned}$$

Therefore μ is invariant under f.

2. Suppose μ is ergodic for f. Let $A \in \mathscr{B}(Y)$ such that $g^{-1}(A) = A$. Since h is the conjugacy between f and g, $h^{-1} \circ g^{-1}(A) = f^{-1} \circ h^{-1}(A)$. Therefore $f^{-1} \circ h^{-1}(A) = h^{-1}(A)$. Since μ is ergodic for f, $\mu(h^{-1}(A)) = 0$ or 1. Then $h_*\mu(A) = \mu(h^{-1}(A)) = 0$ or 1. Hence $\mu_* h$ is ergodic for g.

Now suppose $h_*\mu$ is ergodic for g. Let $A \in \mathscr{B}(X)$ such that $f^{-1}(A) = A$. Then by conjugacy h, $h = g^{-1} \circ h \circ f$. Thus $g^{-1} \circ h(A) = h(A)$. Since $h_*\mu$ is ergodic for g, $h_*\mu(h(A)) = 0$ or 1. Therefore $\mu(A) = \mu(h^{-1} \circ h(A)) = h_*\mu(h(A)) = 0$ or 1. Hence μ is ergodic for f.

□

Now we prove that a map can have uncountable number of invariant, ergodic measures by considering function $f(x) = 2x (\mod 1)$.

Proposition 3.1. *The map $f : [0,1] \to [0,1]$ given by $f(x) = 2x$ mod 1 admits an uncountable family of non_atomic, mutually singular, ergodic, invariant measures.*

Proof. Let $p \in (0,1)$. Consider the map f_p given by

$$f_p(x) = \begin{cases} \frac{x}{p} & \text{if } x \in [0,p) \\ \frac{x}{1-p} - \frac{p}{1-p} & \text{if } x \in [p,1) \end{cases} \tag{3.1}$$

Then there exists a topologically conjugacy h_p, between f_p and f as described in example 1.3. Since h_p is a homeomorphism, h_p and h_p^{-1} are continuous. Therefore they are measurable. Thus h_p is a measurable conjugate between f_p and f. Hence $\mu_p = h_p * m$ is a probability measure on $\mathcal{M}([0,1])$.

- μ_p is invariant and ergodic for f.
  Since m is invariant and ergodic for f_p, by theorem 3.1, μ_p is invariant and ergodic for f.

- μ_p is non_atomic

Let $\{a\} \in \mathcal{M}([0,1])$. Then $\mu_p(\{a\}) = h_p * m(\{a\}) = m(h_p^{-1}(\{a\}))$.
Since h_p is a bijection $h_p^{-1}(\{a\})$ is a singleton set. Then $m(h_p^{-1}(\{a\})) = 0$ as m is non_atomic. Therefore $\mu_p(\{a\}) = 0$. Hence μ_p is non_atomic.

- μ_p and μ_q are mutually singular measures for $p, q \in (0,1)$ such that $p \neq q$.

Let $E_p = \left\{ x \in [0,1] \,\middle|\, \lim_{n\to\infty} \frac{1}{n} \# \left\{ 1 \leq i \leq n | f^i(x) \in \left[0, \frac{1}{2}\right) \right\} = p \right\}$ and

$$E_q = \left\{ x \in [0,1] \,\middle|\, \lim_{n\to\infty} \frac{1}{n} \# \left\{ 1 \leq i \leq n | f^i(x) \in \left[0, \frac{1}{2}\right) \right\} = q \right\}.$$

Then $E_p \cap E_p = \emptyset$. Also $\mu_p(E_p) = 1$ and $\mu_q(E_q) = 1$. To show that, let

$$A_p = \left\{ x \in [0,1] \,\middle|\, \lim_{n\to\infty} \frac{1}{n} \# \left\{ 1 \leq i \leq n | f_p^i(x) \in [0,p) \right\} = p \right\}.$$

Then by Birkhoff's Ergodic theorem, $m(A_p) = 1$. Since h_p is measurable conjugacy between $f+p$ and f, $E_p = h_p(A_p)$. Hence $\mu_p(E_p) = 1$. Similarly we can show $\mu_q(E_q) = 1$.

Suppose $\mu_q(E_p) > 0$.
Then $1 = \mu_q([0,1)) \geq \mu_q(E_p \bigcup E_q) = \mu_q(E_p) + \mu_q(E_q) = \mu_q(E_p) + 1 > 1$. This contradicts that μ_q is a probability measure. Hence $\mu_q(E_p) = 0$. Therefore μ_p and μ_q are mutually singular measures for each distinct $p, q \in (0,1)$. i.e for each distinct $p, q \in (0,1)$, μ_p and μ_q are distinct.
Therefore, there are uncountably many non_atomic, mutually singular, ergodic and invariant measures for f.

□

4 Existence of lebesgue absolutely continuous measure

As we discussed in the first chapter, a point x in the space has well described asymptotic distribution if it belongs to $\mathscr{B}_\mu$. Our aim is to describe a large set of points by that measure, i.e. $m(\mathscr{B}_\mu) > 0$. In this chapter we show that full branch maps with bounded distortion, has such a measure.

In the previous chapter, we showed that the full branch map $f(x) = 2x(\mod 1)$ has uncountable invariant, ergodic measures. Here we show that there exists only one ergodic, invariant measure for full branch maps which is absolutely continuous w.r.t lebesgue measure, m.
Following example shows that for a continuous, non full branch map, such a measure can be not unique.

Example 4.1.

Let

$$f(x) = \begin{cases} \frac{1}{2} - 2x & \text{if } x \in [0, \frac{1}{4}] \\ 2x - \frac{1}{2} & \text{if } x \in [\frac{1}{4}, \frac{3}{4}] \\ \frac{5}{2} - 2x & \text{if } x \in [\frac{3}{4}, 1] \end{cases} \tag{4.1}$$

The function f is full branch and affine on $[0, \frac{1}{2}]$. Therefore lebesgue measure, m is invariant and ergodic for f on $[0, \frac{1}{2}]$. Thus $m_{[0,\frac{1}{2}]}$ is invariant and ergodic for f on $[0, 1]$. Also $m_{[0,\frac{1}{2}]} \ll m$. Similarly $m_{[\frac{1}{2},1]}$ is invariant, ergodic and $m_{[\frac{1}{2},1]} \ll m$ for f on $[0, 1]$.

Now let's draw attention to some important observations which will be helpful to prove our main result.

Lemma 4.1. If $f : I \to J$ is a diffeomorphism, where I, J are intervals in $\mathbb{R}$, then for any $B \in \mathscr{B}(J)$ if $m(B) > 0$ then $m(f(B)) > 0$.

Proof. Let $\{I_k\}_{k\in\mathbb{N}} \subseteq I$ such that $B \subseteq \bigcup_{k\in\mathbb{N}} I_k$. Since f^{-1} is differentiable, $f'(x) \neq 0$ for each $x \in I$. Then there exists $\epsilon > 0$ for each $x \in I_k$ such that $|f'(x)| \geq \epsilon > 0$. Then by mean value theoren $\frac{m\left(f(I_k)\right)}{m(I_k)} \geq \epsilon$. Thus $m\left(f(I_k)\right) \geq \epsilon m(I_k)$. Then,

$$\begin{aligned} m\left(f(B)\right) &= \inf\left\{\sum_{k\in\mathbb{N}} m(J_k) : f(B) \subseteq \bigcup_{k\in\mathbb{N}} J_k\right\} \\ &= \inf\left\{\sum_{k\in\mathbb{N}} m(J_k) : B \subseteq \bigcup_{k\in\mathbb{N}} f^{-1}(J_k)\right\} \end{aligned}.$$

Let $I_k = f^{-1}(J_k)$. Then,

$$
\begin{aligned}
\inf\left\{\sum_{k\in\mathbb{N}} m(J_k) : B \subseteq \bigcup_{k\in\mathbb{N}} f^{-1}(J_k)\right\} &= \inf\left\{\sum_{k\in\mathbb{N}} m\left(f(I_k)\right) : B \subseteq \bigcup_{k\in\mathbb{N}} I_k\right\} \\
&\geq \inf\left\{\sum_{k\in\mathbb{N}} \epsilon m(I_k) : B \subseteq \bigcup_{k\in\mathbb{N}} I_k\right\} \\
&= \epsilon \inf\left\{\sum_{k\in\mathbb{N}} m(I_k) : B \subseteq \bigcup_{k\in\mathbb{N}} I_k\right\} \\
&= \epsilon m(B)
\end{aligned}
.
$$

Thus $m\left(f(B)\right) \geq \epsilon m(B)$. Hence $m\left(f(B)\right) > o$ as $m(B) > 0$. □

If f is not a diffeomorphism, it is not necessary to be non-singular. i.e. image of a positive measure set can have zero measure. For example consider the topological conjugacy h_p, in example 1.3 between functions f_p and f.

Example 4.2.

Let $A = \left\{x \in I : \lim_{n\to\infty} \frac{1}{n}\#\{1 \leq j \leq n : f_p^j(x) \in (0,p)\} = p\right\}$.

Then Birkhoff's Ergodic theorem says that $m(A) = 1$. Now let $x \in A$.
Then $\lim_{n\to\infty} \#\frac{1}{n}\left\{1 \leq j \leq n : f^j\left(h(x)\right) \in (0, \frac{1}{2})\right\} = p$, since x and $h(x)$ have same symbolic dynamics. But by Birkhoff's Ergodic theorem,
$\lim_{n\to\infty} \#\frac{1}{n}\left\{1 \leq j \leq n : f^j\left(h(x)\right) \in (0, \frac{1}{2})\right\} = \frac{1}{2}$ a.e.(m).
Thus $m\left\{x \in I : \lim_{n\to\infty} \frac{1}{n}\#\{1 \leq j \leq n : f^j(x) \in (0, \frac{1}{2})\} \neq \frac{1}{2}\right\} = 0$.
Therefore if $p \neq \frac{1}{2}$, $m\left(h(A)\right) = 0$.

Let $\mathscr{P}^{(n)}$ be the partition of full branch map $f^n : I \to I$.

Lemma 4.2. *Let $f : I \to I$ be a full branch map such that for each $n \geq 1$, $w \in \mathscr{P}^{(n)}$ and for all $x, y \in w$; $\log\left|\frac{Df^n(x)}{Df^n(y)}\right| \leq \widetilde{K}\left|Df^n(x) - Df^n(y)\right|$ for some constant $\widetilde{K} > 0$. Let $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} f_*^i m$. Then,*

1. *$\mu_n \ll m$*
2. *if $H_n = \frac{d\mu_n}{dm}$ and $S_i(x) = \sum_{\widetilde{x}=f^{-i}(x)} \frac{1}{|Df^i(\widetilde{x})|}$ then $H_n(x) = \frac{1}{n}\sum_{i=0}^{n-1} S_i(x)$*
3. *there exists $K > 0$ such that $0 < \inf_{n,x} S_n(x) \leq \sup_{n,x} S_n(x) \leq K$ for all $n \geq 1$ and for all $x, y \in I$.*
4. *there exists $K > 0$ such that $0 < \inf_{n,x} H_n(x) \leq \sup_{n,x} H_n(x) \leq K$ for all $n \geq 1$ and for all $x, y \in I$.*

5. for each $n \geq 1$ and $x, y \in I$, $|S_n(x) - S_n(y)| \leq K |S_n(x)| d(x,y) \leq K^2 d(x,y)$

6. for each $n \geq 1$ and $x, y \in I$, $|H_n(x) - H_n(y)| \leq K |H_n(x)| d(x,y) \leq K^2 d(x,y)$.

Proof. 1. Let $A \in \mathscr{B}(I)$ such that $m(A) = 0$.
Then by lemma 4.1, $m\left(f^{-1}(A) \cap w\right) = 0$ for each $w \in \mathscr{P}$. Since i is countable $m\left(f^{-1}(A)\right) = \sum_{w \in \mathscr{P}} m\left(f^{-1}(A) \cap w\right) = 0$. Then by induction $f_*^i m(A) = 0$ for $i = 0, 1, 2, \cdots, n-1$. Therefore $\mu_n(A) = \frac{1}{m} \sum_{i=0}^{n-1} f_*^i m(A) = 0$ for each $n \in \mathbb{N}$. Therefore $\mu_n \ll m$ for each $n \in \mathbb{N}$.

2. Since $\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} f_*^i m$ and H_n is the density of μ_n, $S_i(x)$ is the density of $f_*^i m$ for $i \in \mathbb{N}$. For that let $A \in \mathscr{B}(I)$. For $i \in \mathbb{N}$ and $w \in \mathscr{P}^{(i)}$; $f^i(w) = I$. And since f is a diffeomorphism, $\left(f^{-1}\right)'(x) = \frac{1}{f'\left(f^{-1}(x)\right)}$.
Therefore $m\left(f^{-i}(A) \cap w\right) = \int_A \frac{1}{Df^i\left(f^{-i}(x) \cap w\right)} dm$. Then,

$$\begin{aligned}
\int_A S_i dm &= \int_A \sum_{\widetilde{x} = f^{-i}(x)} \frac{1}{|Df^i(\widetilde{x})|} \\
&= \int_A \sum_{w \in \mathscr{P}^{(i)}} \frac{1}{|Df^i(f^{-i}(x) \cap w)|} dm \\
&= \sum_{w \in \mathscr{P}^{(i)}} \int_A \frac{1}{|Df^i(f^{-i}(x) \cap w)|} dm \quad . \\
&= \sum_{w \in \mathscr{P}^{(i)}} m\left(f^{-i}(A) \cap w\right) \\
&= m\left(f^{-i}(A)\right) \\
&= f_*^i m(A)
\end{aligned}$$

Thus $\int_A H_n dm = \mu_n(A) = \frac{1}{n} \sum_{i=0}^{n-1} f_*^i m(A) = \frac{1}{n} \sum_{i=0}^{n-1} \int_A S_i dm = \int_A \frac{1}{n} \sum_{i=0}^{n-1} S_i dm.$
Therefore $H_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} S_i(x)$ for $x \in I$.

3. Since $f^n : w \to I$ is a diffeomorphism with uniformly bounded distortion for each $n \in \mathbb{N}$, $w \in \mathscr{P}^{(n)}$ and $x, y \in w$, $\left|\frac{Df^n(x)}{Df^n(y)}\right| \leq K$.

Thus for any $x, y \in w$, $|Df^n(x)| \approx |Df^n(y)|$. By mean value theorem there exists $\zeta \in w$ such that $|Df^n(\zeta)| = \left|\frac{I}{w}\right|$. Therefore $|Df^n(\zeta)| = \frac{1}{|w|}$. Thus for $x \in w$, $|Df^n(x)| \approx \frac{1}{|w|}$ as $|Df^n(x)| \approx |Df^n(\zeta)|$.
Therefore $S_n(x) = \sum_{\widetilde{x} = f^{-n}(x)} \frac{1}{|Df^n(\widetilde{x})|} \approx \sum_{w \in \mathscr{P}^{(n)}} |w| = 1$. Hence $S_n(x) \approx 1$.

Thus there exists $K > 0$ such that $\frac{1}{K} \geq S_n(x) \geq K$ for each $n \in \mathbb{N}$ and for each $x \in I$. Therefore $\inf_{n,x} S_n(x) \leq \sup_{n,x} S_n(x) \leq K$.

4. By (3), $\inf_{n,x} H_n(x) = \inf_{n,x} \frac{1}{n} \sum_{i=0}^{n-1} S_i(x) = \frac{1}{n} \sum_{i=0}^{n-1} \inf_{i,x} S_i(x) > 0$ and

$$\sup_{n,x} H_n(x) = \sup_{n,x} \frac{1}{n} \sum_{i=0}^{n-1} S_i(x) = \frac{1}{n} \sum_{i=0}^{n-1} \sup_{i,x} S_i(x) \leq K.$$

Thus $0 < \inf_{n,x} H_n(x) \leq \sup_{n,x} H_n(x) \leq K$.

5. Since $\log \left| \frac{Df^n(x)}{Df^n(y)} \right| < \widetilde{K} \left| f^n(x) - f^n(y) \right|$,

$$\left| \frac{Df^n(x)}{Df^n(y)} \right| \leq \exp \widetilde{K} \left| f^n(x) - f^n(y) \right| \leq 1 + K_1 \left| f^n(x) - f^n(y) \right| \text{ for some } K_1 > 0.$$

Thus $\left| \frac{Df^n(x)}{Df^n(y)} \right| \geq \frac{1}{1 + K_1 \left| f^n(x) - f^n(y) \right|} \geq 1 - K_2 \left| f^n(x) - f^n(y) \right|$ for some $K_2 > 0$.

Therefore $-K_2 \left| f^n(x) - f^n(y) \right| \leq \left| \frac{Df^n(x)}{Df^n(y)} \right| - 1 \leq K_1 \left| f^n(x) - f^n(y) \right|$.

Let $K_3 = \max\{K_1, K_2\}$. Then $\left| \left| \frac{Df^n(x)}{Df^n(y)} \right| - 1 \right| \leq K_3 \left| f^n(x) - f^n(y) \right|$ for $x, y \in I$ and $n \in \mathbb{N}$.

Since the number of subintervals w in $\mathscr{P}^n$ is countable, let $f^n(\widetilde{x_i}) = x$ and $f^n(\widetilde{y_i}) = y$ for each $i \in \mathbb{N}$. Then,

$$\begin{aligned}
|S_n(x) - S_n(y)| &= \left| \sum_{\widetilde{x} = f^{-n}(x)} \frac{1}{|Df^n(\widetilde{x})|} - \sum_{\widetilde{y} = f^{-n}(y)} \frac{1}{|Df^n(\widetilde{y})|} \right| \\
&= \left| \sum_{i \in \mathbb{N}} \left(\frac{1}{|Df^n(\widetilde{x_i})|} - \frac{1}{|Df^n(\widetilde{y_i})|} \right) \right| \\
&\leq \sum_{i \in \mathbb{N}} \left| \frac{1}{|Df^n(\widetilde{x_i})|} - \frac{1}{|Df^n(\widetilde{y_i})|} \right| \\
&= \sum_{i \in \mathbb{N}} \frac{1}{|Df^n(\widetilde{x_i})|} \left[1 - \left| \frac{Df^n(\widetilde{x_i})}{Df^n(\widetilde{y_i})} \right| \right] \\
&\leq \sum_{i \in \mathbb{N}} \frac{1}{|Df^n(\widetilde{x_i})|} K_3 \left| f^n(\widetilde{x_i}) - f^n(\widetilde{y_i}) \right| \\
&= K_3 |x - y| \sum_{i \in \mathbb{N}} \frac{1}{|Df^n(\widetilde{x_i})|} \\
&= K_3 S_n(x) |x - y|
\end{aligned}$$

6.

$$\begin{aligned}
|H_n(x) - H_n(y)| &= \left| \frac{1}{n}\sum_{i=0}^{n-1} S_i(x) - \frac{1}{n}\sum_{i=0}^{n-1} S_i(y) \right| \\
&\leq \frac{1}{n}\sum_{i=0}^{n-1} |S_i(x) - S_i(y)|
\end{aligned}$$

Then by lemma 4.2 - (5), $|S_i(x) - S_i(y)| \leq K_3 S_i(x)\,|x-y|$ for each $i \in \mathbb{N}$. Therefore,

$$\begin{aligned}
\frac{1}{n}\sum_{i=0}^{n-1} |S_i(x) - S_i(y)| &\leq \frac{1}{n}\sum_{i=0}^{n-1} K_3 S_i(x)\,|x-y| \\
&= K_3 H_n(x)\,|x-y| \\
&\leq K K_3\,|x-y|
\end{aligned}.$$

□

Theorem 4.1. *Let $f : I \to I$ be a full branch map with $\sup_{n\geq 1}\sup_{w\in\mathscr{P}^{(n)}}\sup_{x,y\in w}\log\left|\frac{Df^n(x)}{Df^n(y)}\right| < \infty$. Then f admits a unique ergodic invariant probability measure μ, which is absolutely continuous w.r.t lebesgue measure.*

Proof. Consider $\mu_n = \frac{1}{m}\sum_{i=0}^{n-1} f_*^i m$. Then by lemma 4.2, $\mu_n \ll m$. Then by Radon Nikodym theorem, there exists a positive borel measurable function H_n such that $\mu_n(A) = \int_A H_n dm$ for $A \in \mathscr{B}(I)$ and for each $n \in \mathbb{N}$. By lemma 4.2, $\{H_n\}$ is a uniformly bounded and equicontinuous sequence of functions. Thus by Arzela Ascoli theorem, $\{H_n\}$ has a convergent subsequence $\{H_{n_k}\}$. Say $\{H_{n_k}\}$ converges to H.
Now let's define $\mu(A) = \int_A H dm$. Since integral of a function over a measure zero set is zero, $\mu \ll m$.

Since f satisfy $\sup_{n\geq 1}\sup_{w\in\mathscr{P}^{(n)}}\sup_{x,y\in w}\log\left|\frac{Df^n(x)}{Df^n(y)}\right| < \infty$, m is ergodic for f. Then μ is ergodic by lemma 1.1 as $\mu \ll m$. To prove that μ is invariant for f, let $A \in \mathscr{B}(I)$. Consider,

$$\begin{aligned}
\mu(A) - \mu\left(f^{-1}(A)\right) &= \int_A H dm - \int_{f^{-1}(A)} H dm \\
&= \int_A \lim_{n_k\to\infty} H_{n_k} dm - \int_{f^{-1}(A)} \lim_{n_k\to\infty} H_{n_k} dm \\
&= \lim_{n_k\to\infty}\left(\int_A H_{n_k} dm - \int_{f^{-1}(A)} H_{n_k} dm\right) \\
&= \lim_{n_k\to\infty}\left(\mu_{n_k}(A) - \mu_{n_k}(f^{-1}(A))\right) \\
&= \lim_{n_k\to\infty}\left(\frac{1}{n_k}\sum_{i=0}^{n_k-1} f_*^i m(A) - \frac{1}{n_k}\sum_{i=0}^{n_k-1} f_*^i m\left(f^{-1}(A)\right)\right) \\
&= \lim_{n_k\to\infty}\left(\frac{1}{n_k}\sum_{i=0}^{n_k-1} m\left(f^{-i}(A)\right) - \frac{1}{n_k}\sum_{i=1}^{n_k} m\left(f^{-i}(A)\right)\right) \\
&= \lim_{n_k\to\infty}\left(\frac{1}{n_k} m(A) - \frac{1}{n_k} m\left(f^{-n_k}(A)\right)\right) \\
&= 0
\end{aligned}$$

Therefore μ is f - invariant, ergodic and absolutely continuous w.r.t. lebesgue measure, m.

To prove uniqueness, let's assume μ_1 is a probability measure such that ergodic, f-invariant and absolutely continuous w.r.t m.
Let $\varphi \in C^\circ(I)$ and $A = \left\{ x \in I : \lim_{n\to\infty} \sum_{k=0}^{n-1} \varphi \circ T^k(x) = \int f d\mu_1, \right\}$. Then by Birkhoff's Ergodic Theorem $\mu_1(A) = 1$. Then $m(A) > 0$ as $\mu_1 \ll m$. Since $\inf H_n > 0$; $H > 0$. Thus $\mu(B) > 0$ for any $B \in \mathscr{B}(I)$ such that $m(B) > 0$ by the definition of $\mu(B)$.

Therefore $\mu(A) > 0$. Hence $A_1 = \left\{ x \in A : \lim_{n\to\infty} \sum_{k=0}^{n-1} \varphi \circ T^k(x) = \int \varphi d\mu \right\} \neq \emptyset$. Let $x \in A_1$. Then $\int \varphi d\mu_1 = \int f d\mu$ for $\varphi \in C^\circ(I)$. Then by theorem 1.3, $\mu_1 = \mu$.

Therefore, for full branch maps; invariant, ergodic and absolutely contimuous measure w.r.t lebesgue measure is unique.

□

Bibliography

[1] Stefano, L., 2013, "Introduction to Smooth Ergodic Theory," *Lecture notes*, The Abdus Sakan International Center for Theoretical Physics , Italy.

[2] Karl, P. , 1983, "Ergodic Theory," *Cambridge studies in advanced mathematics 2*, Vol .2, Cambridge University Press, Cambridge.

[3] Peter W. , 1979, "An Introduction to Ergodic Theory," *Graduate Texts in Mathematics*, Springer-Verlag, New York.

[4] Royden, H.L., 1988, "Real Analysis," , Pearson Education , 3rd ed.

[5] Ricardo M. , 1983, "Ergodic Theory and Differentiable Dynamics," *A series of modern surveys in Mathematics*, Springer-Verlag , Berlin.