UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

INSTITUTO DE MATEMÁTICA

CURSO DE MESTRADO EM MATEMÁTICA

GIOVANNY ANDRÉS FUENTES SALVO

DYNAMICS SYSTEMS CHARACTERIZATION THROUGH THE FROBENIUS-PERRON OPERATOR AND OTHERS MARKOV OPERATORS: SOME APPLICATION TO SECOND LAW OF THERMODYNAMICS

RIO DE JANEIRO

2020

GIOVANNY ANDRÉS FUENTES SALVO

DYNAMICS SYSTEMS CHARACTERIZATION THROUGH THE FROBENIUS-PERRON OPERATOR AND OTHERS MARKOV OPERATORS: SOME APPLICATION TO SECOND LAW OF THERMODYNAMICS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Matemática.

Orientador: Prof. Maurizio Monge, Ph.D

RIO DE JANEIRO

2020

Fuentes Salvo, Giovanny Andrés

Dynamics systems characterization through the
Frobenius Perron operator and others
Markov operators: some application to
second law of thermodynamics
on/ Giovanny Andrés Fuentes Salvo. -Rio de Janeiro
: IM/UFRJ,2019.

VII,50 p. 29.27cm.

Orientador: Maurizio Monge Ph.D

Dissertação (mestrado)-IM/UFRJ/Programa de Matemática,2019.

Referências Bibliográficas: p. 50.

1. aaaaa. 2. ssss.3. Hddddd. 4. dddddd.I.D.Sc.,
Adán José Corcho. II. universidade Federal do Rio
de Janeiro, IM, Programa de Matemática. III. Título.

GIOVANNY ANDRÉS FUENTES SALVO

DYNAMICS SYSTEMS CHARACTERIZATION THROUGH THE FROBENIUS-PERRON OPERATOR AND OTHERS MARKOV OPERATORS: SOME APPLICATION TO SECOND LAW OF THERMODYNAMICS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Matemática.

Aprovado em ______ de ________________ de _________ .

BANCA EXAMINADORA:

Prof. Maurizio Monge, Ph.D

Prof. Nisoli Isaia, Ph.D

Prof. Sergio Romaña, Ph.D

Prof. Alejandro Kocsard, Ph.D.

I dedicate this work to my beloved
parents who were always for me

AGRADECIMENTOS

Epígrafe: É um item onde o autor
apresenta a citação de um texto que seja
relacionado com o tema do trabalho,
seguido da indicação de autoria do
mesmo. (texto iniciando do meio da
página alinhado à direita)

Nome do autor da epígrafe

RESUMO

O objetivo deste trabalho é caracterizar os sistemas dinâmicos, ou seja, se eles são ergódicos, mistos ou exatos, através do operador Frobenius Perron. Vamos também fazer uma conexão com a segunda lei da termodinâmica

Palavras-chave: Palavrachave1. Palavrachave2. Palavrachave3.

ABSTRACT

The objective of this work is to characterize the dynamic systems, that is, if they are ergodic, mixing or exact, through the Frobenius Perron operator. Also make a connection with the second law of thermodynamics

Keywords: Keyword1. Keyword2. Keyword3.

LISTA DE FIGURAS

SUMÁRIO

1 INTRODUCTION

In this work, we are going to study an important Markov operator in the study of the dynamics of density functions, the Frobenius-Perron operator. This operator allows us to study the evolution of the density functions under a dynamical system. This operator takes a density function in an arbitrary system state and gives us a new density function, according to the system's dynamics . For example if the system tends to accumulate all the orbits at a point in space, then the operator of Frobenius-Perron, close to makes all the densities evolve to a density such that it has a maximum that point, since it is at that point we should have the greatest probability of finding particles in a sufficiently large time. In the case that the system tends to disperse all orbits throughout the space evenly, then the Frobenius-Perron operator will make all densities at a constant density. On the other hand, knowing how densities evolve gives us information about the behavior of all the orbits of the system.

Next we will give a brief summary of what we will see in each chapte.

In chapter one, we will introduce Markov's operators, especially the Frobenius-Perron operator and its adjoint operator, the Koopman operator.

In chapter two, we will consider three types of dynamical systems; the ergodicones, the mixting ones and the exact ones. In the main theorem of this chapter shows the equivalence of the three types of systems mentioned above with the convergence modes of the Frobenius-Perron operator sequence.

In chapter three, we will study sufficient conditions for the existence of fixed points for Markov operators, we will call these fixed points stationary densities.

In the chapter chapter, we will study the behavior of the entropy of Bolzmann-Gibbs in relation to the dynamic behavior of the Markov operators, in particular when the Markov operator is the Frobenius-Perron operator. We end with a theorem that establishes a condition for the accuracy of the system, in relation to the entropy of Bolzmann-Gibbs.

In chapter five, we will give an interpretation of the results seen in the present work in the context of thermodynamics, in particular with the second law of thermodynamics.

2 THE FROBENIUS-PERRON OPERATOR

In this chapter we introduce Markov operator, with the objective of study the dynamics of a density within a dynamical system or to characterize the associated dynamical system itself, the Frobenius-Perron operator, also known as the transfer operator.

2.1 THE MARKOV OPERATOR

We will start by defining what is a Markov operator.

Definition 1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space. Any linear operator $P : L^1 \to L^1$ satisfying for all $f \in L^1$ such that $f \geq 0$:*

1. *$Pf \geq 0$,*
2. *$\|Pf\| = \|f\|$*

is called a Markov Operator

The Markov Operator have the following property if $f, g \in L^1$ then $Pf(x) \geq Pg(x)$ whenver $f(x) \geq g(x)$. Indeed $(f(x) - g(x)) \geq 0$ implies that $P(f(x) - g(x)) \geq 0$ and by items 1 of the Definition 2.1 of P.

To demonstrate further inequalities, we offer the following Proposition:

Proposition 2 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space and P is Markov operator, then for every $f \in L^1$:*

1. *$(Pf(x))^+ \leq P(f(x)^+)$,*
2. *$(Pf(x))^- \leq P(f(x)^-)$,*
3. *$|Pf(x)| \leq P|f(x)|$,*

4. $\|Pf(x)\| \leq \|f(x)\|$.

Proof Note that from Definition of f^+ and f^-, it following that:

$$(Pf)^+ = (Pf^+ - Pf^-)^+ = Max\{0, Pf^+ - Pf^-\} \leq Max\{0, Pf^+\} = Pf^+.$$

Analogous is the proof for Pf^-. The inequality 3 following from 1 and 2, namely:

$$|Pf| = (Pf)^+ + (PF)^- \leq P(f^+) + P(F)^- = P(f^+ = Pf^-) = P|f| \ .$$

Finally, by integrating 3 over $\mathcal{X}$, we obtain:

$$\|Pf\| = \int_{\mathcal{X}} |Pf(x)|d\mu \leq \int_{\mathcal{X}} P|f(x)|d\mu = \|f\|,$$

which confirms 4.

The inequality 4 is very important, any operator P that satisfies it is called a contraction.

Definition 3 *Support of function g we simply mean the set of all x such that $g(x) \neq 0$, thar is:*

$$Supp(g) = \{x; g(x) \neq 0\} \ .$$

This is not the customary Definition of the support of a function, which is $closure\{x; g(x) \neq 0\}$, but, because the customary Definition requires the introduction of topological notions not used elsewhere.

Proposition 4 *Let P is Markov operator, then $\|Pf(x)\| = P\|f(x)\|$ if and only if $P(f(x)^+)$ and $P(f(x)^-)$ have disjoint supports.*

Proof: We start from the inequality:

$$|Pf^+ - Pf^-| \leq |Pf^+| + |Pf^-|.$$

Clearly the inequality will be strong if both $Pf^+ > 0$ and $Pf^- > 0$, while the inequality holds if $Pf^+ = 0$ or $Pf^- = 0$. thus, by integrating over the space $\mathcal{X}$, we obtain:

$$\int_{\mathcal{X}} |Pf^+ - Pf^-| d\mu = \int_{\mathcal{X}} |Pf^+| d\mu + \int_{\mathcal{X}} |Pf^-| d\mu$$

if and only if there is not $A \in \mathcal{A}$ with $\mu(A) > 0$ such that $Pf^+ > 0$ and $Pf^- > 0$ for $x \in A$, that is Pf^+ and Pf^- have disjoint support. Since $f = f^+ - f^-$, the left-hand integral is simply $\|Pf\|$. Further, the right-hand side is $\|Pf^+\| + \|Pf^-\| = \|f^+\| + \|f^+\| = \|f\|$, so the Proposition is proved.

Definition 5 *If P is Markov operator and, for some $f \in L^1$ we have that $Pf = f$, then f is called a fixed point of P.*

Proposition 6 *Let P be a Markov operator and $f \in L^1$, then if $Pf = f$, then $Pf^+ = f^+$ and $Pf^- = f^-$.*

Proof: Note that from $Pf = f$, we obtain:

$$f^+ = (Pf)^+ \leq Pf^+ \text{ and } f^- = (Pf)^- \leq Pf^-$$

Hence

$$\begin{aligned} \int_{\mathcal{X}} (Pf^+ - f^+) d\mu + \int_{\mathcal{X}} (Pf^- - f^-) d\mu &= \int_{\mathcal{X}} (Pf^+ - Pf^-) d\mu - \int_{\mathcal{X}} (f^+ - f^-) d\mu \\ &= \int_{\mathcal{X}} Pf d\mu + \int_{\mathcal{X}} f d\mu = \|P|f|\| - \|f\| \end{aligned}$$

however, by the contractive property of P we know that:

$$\|P|f|\| - \|f\| \leq 0.$$

Since both the integrands $(Pf^+ - f^+)$ and $(Pf^- - f^-)$ are non-negative thus last inequality is possible only if:

$$Pf^+ = f^+ \text{ and } Pf^- = f^-$$

Definition 7 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be measure space and the set*

$$D(\mathcal{X}, \mathcal{A}, \mu) = \{f \in L^1(\mathcal{X}, \mathcal{A}, \mu); f \geq 0, \|f\| = 1\}.$$

Any function $f \in D(\mathcal{X}, \mathcal{A}, \mu)$ *is called a* ***density****.*

Definition 8 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *a measure space, if* $f \in L^1(\mathcal{X}, \mathcal{A}, \mu)$ *and* $f \geq 0$ *then the measur said to be absolutely continuous with respect to* μ *if can be written like:*

$$\mu_f(A) = \int_A f(x) d\mu$$

and f *is called the* ***Radon-Nikodym derivative*** *of* μ_f *with respect to* μ*.*

Definition 9 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be measure space and* P *be a Markov operator. Any* $f \in D$ *that satisfies* $Pf = f$ *is called a stationary density of* P*.*

2.2 THE FROBENIUS-PERRON OPERATOR

Before formally defining the Frobenius-Perron operator of a dynamical system, let's see an intuitive construction proposed by Gora and Boyasky in [7]

Let us take a dynamical system $S : \mathcal{X} \to \mathcal{X}$ and A a subset of $\mathcal{X}$, now consider a random variable x and a function of density f, we can calculate the probability that the random variable x belongs to A, simply calculating the integral.

$$\int_A f d\mu$$

where μ is the normalized measure of lebesgue in $\mathcal{X}$.

Let's imagine we need to know, what is the probability that the random variable x belongs to $S(A)$? an innocent answer would be to calculate it in the same way using the same probability density, but we have bad news, the density in most of the times will also be affected by the dynamics of the dynamical system, then the problem is reduced to determine (if it is possible) the new state of the density function. Then we try to deduce what is the new state of the density function or at least to determine an equation (as a function of the initial density f and the dynamical system S), that allows us to define and guarantee their existence and unity.

Let's start by trying to calculate the probability that $S(x)$ belongs to A.

$$Prob\{S(x) \in A\} = Prob\{x \in S^{-1}(A)\} = \int_{S^{-1}(A)} f d\mu$$

To obtain a probability density function for $S(x)$, we have to write this last integral as

$$\int_A \varphi d\mu$$

for some function φ. let's say the measure of the probability that $S(x) \in A$ is $\lambda(A)$, then

$$\lambda(A) = \int_{S^{-1}(A)} f d\mu$$

Only up to this point, we can arrive with the few demands that we have asked S and f. We need to add some more condition, to be able to use some theorem, that guarantees us the existence and uniqueness of the evolution of our density function. Let us test with the following condition *that the measure (all about μ) of the inverse image of a set with null measure is null.* Let's analyze what consequences this condition would have.

If $\mu(A) = 0$ this implies that $\mu(S^{-1}(A)) = 0$ which in turn implies that $\lambda(A) = 0$. In the other words $\lambda << \mu$ Then, by the Radon-Nikodym Theorem, there exists a $\varphi \in L^1$ such that for all measurable sets A such that

$$\lambda(A) = \int_A \varphi d\mu$$

and φ is unique a.e., and depends on S and f. Set $\varphi = Pf$. Thus, the probability density function f has been transformed to a new probability density function Pf. We will call to P the Frobenius-Perron Operator corresponding to S and in addition we will call all function that satisfies the condition that we proposed above of no-sigular. After all this introduction, I will give the formal Definition of the Frobenius-Perron operator

Definition 10 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space. A measurable transformation $S : \mathcal{X} \to \mathcal{X}$ is non-singular if:*

$$\mu(S^{-1}(A)) = 0 \textit{ for all } A \in \mathcal{A} \textit{ such that } \mu(A) = 0.$$

Definition 11 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, if $S : \mathcal{X} \to \mathcal{X}$ is a non-singular transformation the unique operator $P : L^1 \to L^1$ defined by equation:*

$$\int_A Pf(x)d\mu = \int_{S^{-1}(A)} f(x)d\mu \textit{ for } A \in \mathcal{A}$$

is called the ***Frobenius-Perron Operator*** *corresponding to S.*

Proposition 12 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, and $S : \mathcal{X} \to \mathcal{X}$ is a non-singular transformation, let P the associated Frobenius-Perron operator to S, then*

1. $P(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 Pf_1 + \lambda_2 Pf_2$, *for* $f_1, f_2 \in \mathcal{L}^1$ *and* $\lambda_1, \lambda_2 \in \mathbb{R}$
2. $Pf \geq 0$ *if* $f \geq 0$, *for* $f \in \mathcal{L}^1$,
3. $\displaystyle\int_{\mathcal{X}} Pf d\mu = \int_{\mathcal{X}} f d\mu$, *for* $f \in \mathcal{L}^1$,
4. *if* $S_n = \overbrace{S \circ \cdots \circ S}^{n}$ *and P_n is the Frobenius-Perron operator corresponding to S, then $P_n = P^n$, where P is the Frobenius-Perron operator corresponding to S.*

Proof Let $f_1, f_2, f \in L^1$, $A \in \mathcal{A}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, then from the Definition of the Frobenius-Perron operator we obtain:

1. $\displaystyle\int_A P(\lambda_1 f_1 + \lambda_2 f_2)d\mu = \lambda_1 \int_{S^{-1}(A)} f_1 d\mu + \lambda_2 \int_{S^{-1}(A)} (f_2)d\mu$
$\displaystyle= \int_A \lambda_1 Pf_1 + \lambda_2 Pf_2 d\mu$. Therefore $P(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 Pf_1 + \lambda_2 Pf_2$.
2. $\displaystyle\int_A Pf d\mu = \int_{S^{-1}(A)} f d\mu \geq 0$. Since $f \geq 0$.
3. Obvious since $S^{-1}(\mathcal{X}) = \mathcal{X}$.

4. $\displaystyle\int_A P_n f d\mu = \int_{S^{-n}(A)} f d\mu = \int_{S^{1-n}(A)} P f d\mu = \cdots = \int_A P^n f d\mu.$

Remark: From points 2 and 3, we can see that the Frobenius-Perron operator is a Markov operator.

Proposition 13 *The Frobenius-Perron operator is a contraction.*

Proof Let $f \in \mathcal{L}^1$. Let $f^+ = \max(f, 0)$ and $f^- = -\min(0, f)$, then $f^+, f^- \in \mathcal{L}^1$, $f = f^+ - f^-$ and $|f| = f^+ - f^-$. By linearity, we have

$$Pf = P(f^+ - f^-) = Pf^+ - Pf^-.$$

Hence,

$$|Pf| \leq |Pf^+| + |Pf^-| = Pf^+ + Pf^- = P|f|$$

and

$$\|Pf\| = \int_{\mathcal{X}} |Pf| d\mu \leq \int_{\mathcal{X}} P|f| d\mu = \int_{\mathcal{X}} |f| d\mu = \|f\|$$

Corollary 13.1 *The Frobenius-Perron operator is an continuous operator.*

In some special case equation of the Definition 11 allows us to obtain an explicit form for Pf. If $\mathcal{X} = [a, b]$ and $A = [a, x]$ whit $a \leq x \leq b$, then

$$\int_a^x Pf(s) ds = \int_{S^{-1}([a,x])} f(s) ds$$

and differentiating

$$Pf(x) = \frac{d}{dx} \int_{S^{-1}([a,x])} f(s) ds.$$

It is important to note in the special case where the transformation S is differentiable and invertible then S must be monotone.

Suppose S is an increasing function and S^{-1} has a continuous derivative, then:

$$S^{-1}([a,x]) = [S^{-1}(a), S^{-1}(x)]$$

and from

$$Pf(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f(s)ds = f(S^{-1}(x))\frac{d}{dx}(S^{-1}(x)).$$

If S is decreasing, then the sing of the right-hand side is reversed. Thus, in the general one-dimensional case, for S differentiable and invertible with continuous $\dfrac{dS^{-1}(x)}{dx}$,

$$Pf(x) = f(S^{-1}(x))\left|\frac{d}{dx}S^{-1}(x)\right|.$$

Example 14 *Let $S : \mathbb{R} \to \mathbb{R}$ given by $S(x) = \exp(x)$, then the associated Frobenius-Perron operator to S is*

$$Pf(x) = \frac{1}{x}f(\log(x)).$$

Consider what happens to an initial f given by

$$f(x) = \frac{1}{2}1_{[-1,1]}$$

under the action of P, the function f is carried into

$$Pf(x) = \frac{1}{2x}1_{[e^{-1},e]}(x)$$

Now applying the operator again, we have

$$P^2f(x) = \frac{1}{2x^2}1_{[e^{e^{-1}},e^e]}(x)$$

Example 15 *Let $S : [0,1] \to [0,1]$ given by $S(x) = 4x(1-x)$ on the measure space $([0,1], \mathcal{B}, \mu)$ where $\mathcal{B}$ is σ-algebra and μ is the Borel measure. First we calculate the inverse image of the interval with x belongs to the interval $[0,1]$*

$$S^{-1}([0,x]) = \left[0, \frac{1}{2} - \frac{1}{2}\sqrt{1-x}\right] \cup \left[\frac{1}{2} + \frac{1}{2}\sqrt{1-x}, 1\right].$$

Now we calculate the Frobenius-Perron operator associated with S

$$Pf(x) = \frac{d}{dx}\int_0^{\frac{1}{2}-\frac{1}{2}\sqrt{1-x}} f(u)du + \frac{d}{dx}\int_{\frac{1}{2}+\frac{1}{2}\sqrt{1-x}}^{1} f(u)du$$

and computing the integral, we have

$$Pf(x) = \frac{1}{4\sqrt{1-x}}\left[f\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-x}\right) + f\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-x}\right)\right]$$

Example 16 *Let* $S : [0,1] \to [0,1]$ *given by* $S(x) = rx \mod 1$ *where* $r > 1$ *is an integer, on the measure space* $([0,1], \mathcal{B}, \mu)$ *where* $\mathcal{B}$ *is* σ*-algebra and* μ *is the Borel measure, then we have* $S^{-1}([0,x])$ *is*

$$\bigcup_{i=0}^{r-1}\left[\frac{i}{r}, \frac{i}{r}+\frac{x}{r}\right]$$

and the Frobenius-Perron operator is

$$Pf(x) = \frac{d}{dx}\sum_{i=0}^{r-1}\int_{\frac{i}{r}}^{\frac{i}{r}+\frac{x}{r}} f(u)du = \frac{1}{r}\sum_{i=0}^{r-1} f\left(\frac{i}{r}+\frac{x}{r}\right)$$

Proposition 17 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be measure space and* $S : \mathcal{X} \to \mathcal{X}$ *be a non-singular transformation and* P *the associated Frobenius-Perron operator to* S*. Assume that an* $f \geq 0$ *and* $f \in L^1$ *is given then:*

$$Supp(f) \subset S^{-1}(Supp(Pf))$$

and, more generally, for every set $A \in \mathcal{A}$ *the following equivalence holds:*

$$Pf = 0 \text{ for } x \in \mathcal{A} \text{ if and only if } f(x) = 0 \text{ for } x \in S^{-1}(A)$$

Proof: By the Definition of the Frobenius-Perron operator, we obtain:

$$\int_A Pf(x)d\mu = \int_{S^{-1}} f(x)d\mu$$

or

$$\int_{\mathcal{X}} 1_A Pf(x)d\mu = \int_{\mathcal{X}} 1_{S^{-1}(A)} f(x)d\mu.$$

Thus $Pf(x) = 0$ on A implies, by property of the Lebesgue integral, that $f(x) = 0$ for $x \in S^{-1}(A)$ and vice versa. Now setting $A = \mathcal{X} \setminus Supp(Pf)$ we obtain $Pf = 0$ for $x \in A$, consequently $f(x) = 0$ for $x \in S^{-1}(A)$. Which means that $Supp(f) \subset \mathcal{X} \setminus S^{-1}(A)$ since $S^{-1}(A) = \mathcal{X} \setminus S^{-1}(Supp(Pf))$, then

$$Supp(f) \subset \mathcal{X} \setminus (\mathcal{X} \setminus S^{-1}(Supp(Pf))) \text{ if and only if } Supp(f) = S^{-1}(Supp(Pf))$$

■

In the case of arbitrary $f \in L^1$, then in Proposition 17 we only have if $f(x) = 0$ for all $x \in S^{-1}(A)$, then $Pf(X) = 0$ for all $x \in A$.

Theorem 18 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation, and $f : \mathcal{X} \to \mathcal{X}$ measurable function such that, $f \circ g \in L^1(\mathcal{X}, \mathcal{A}, \mu)$ then for every $A \in \mathcal{A}$:*

$$\int_{S^{-1}(A)} f(S(x))d\mu = \int_A f(x)d\mu S^{-1} = \int_A f(x)J^{-1}dx.$$

Where μS^{-1} denote the measure:

$$\mu S^{-1}(B) = \mu(S^{-1}(B)) \text{ for } B \in \mathcal{A}.$$

and J^{-1} is the density of μS^{-1} with respect to μ, that is:

$$\mu(S^{-1}(B)) = \int_B J^{-1}dx \text{ for } B \in \mathcal{A}.$$

We use the notation $J^{-1}(x)$ to draw the connection with differentiable invertible trasformation on $\mathbb{R}^d$, in which case $J(x)$ is the determinant of jacobian matrix:

$$J(x) = det\left(\frac{dS(x)}{dx}\right) \text{ and } J^{-1}(x) = det\left(\frac{dS^{-1}(x)}{dx}\right)$$

Proof Of Theorem 18 To prove this change of variables theorem, and fist take:

$$f(x) = 1_B(x)$$

so that $f(S(x)) = 1_B(S(x)) = 1_{S^{-1}(B)}(x)$ and, hence,

$$\begin{aligned}
\int_{S^{-1}(A)} f(S(x))d\mu &= \int_{\mathcal{X}} 1_{S^{-1}(A)} f(s) d\mu \\
&= \int_{\mathcal{X}} 1_{S^{-1}(A)} 1_{S^{-1}(B)} d\mu \\
&= \mu(S^{-1}(A) \cap S^{-1}(B)) \\
&= \mu(S^{-1}(A \cap B))
\end{aligned}$$

The second integral of Theorem may be written as:

$$\begin{aligned}
\int_A f(x) d\mu S^{-1} &= \int_{\mathcal{X}} 1_A 1_B d\mu S^{-1} \\
&= \mu(S^{-1}(A \cap B)).
\end{aligned}$$

Where the third and last integral has the form:

$$\begin{aligned}
\int_A f(x) J^{-1} dx &= \int_A 1_B J^{-1} dx \\
&= \int_{A \cap B} J^{-1} dx \\
&= \mu(S^{-1}(A \cap B)).
\end{aligned}$$

And by linearity of the integral of the Lebesgue, we proved Theorem.

■

Corollary 18.1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ an invertible non-singular transformation (S^{-1} non-singular) and P the associated Frobenius-Perron operator, then for every $f \in L^1$:*

$$Pf(x) = f(S^{-1}(x)) J^{-1}(x).$$

Proof: By the Definition of P, for $A \in \mathcal{A}$ we obtain:

$$\int_A Pf d\mu = \int_{S^{-1}} f(x) d\mu.$$

changer the variable in the right hand integral with $y = S(x)$ so that:

$$\int_{S^{-1}} f(x) d\mu = \int_A f(S^{-1}(y)) J^{-1}(x) d\mu$$

By theorem 18 thus we obtain:

$$\int_A Pf d\mu = \int_A f(S^{-1}(x))J^{-1}(x)d\mu$$

with the result that:

$$Pf(x) = f(S^{-1}(x))J^{-1}(x).$$

■

2.3 THE KOOPMAN OPERATOR

To finish this chapter, we will briefly study the Koopman operator and will limit to analyzing some of its properties. In the following chapters, we will analyze theorems and Propositions related to this operator. For this section we will use the following notation defined in Walter Rudin's functional analysis book. It will be convenient to designate elements of the dual space $\mathcal{X}^*$ of $\mathcal{X}$ by x^* and to write $\langle x, x^* \rangle$ in place of $x^*(x)$.

Theorem 19 *Let $\mathcal{X}$ and $\mathcal{Y}$ normed space. To each $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ corresponds a unique $T^* \in \mathcal{B}(\mathcal{Y}^*, \mathcal{X}^*)$ that satisfies*

$$\langle Tx, y^* \rangle = \langle x, T^*y^* \rangle$$

for all $x \in \mathcal{X}$ and all $y^* \in \mathcal{Y}^*$. Moreover T^* satisfies

$$\| T^* \| = \| T \|$$

In several books of functional analysis define the dual of a linear transformation in a Banach space using the following notation

$$y^*(Tx) = T^*y^*(x)$$

Taking advantage of Rudin's notation, we can define with abuse of notation, the scalar product between functions and functions of its dual

Definition 20 *Let $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^{p^*}$, we define the scalar product of two functions by*

$$\langle f, g \rangle = \int_{\mathcal{X}} fg d\mu.$$

in usual notation this Definition would be equivalent to saying

$$g(f) := \int_{\mathcal{X}} fg d\mu = \langle f, g \rangle$$

An important relation we often use is the Cauchy-Hölder inequality. Thus, if $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^{p^*}$, then

$$|\langle f, g \rangle| \leq \|f\|_{\mathcal{L}^p} \|g\|_{\mathcal{L}^{P^*}}$$

We will briefly study the Koopman operator and will limit ourselves to analyzing some of its properties.

Definition 21 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation and $f \in L^\infty$, the operator $U : L^\infty \to L^\infty$ defined by $Uf(x) = f(S(x))$ is called the* ***Koopman Operator*** *with respect to S*

The operator U has some important properties:

1. $U(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 U(f_1) + \lambda_2 U(f_2)$ for all $f_1, f_2 \in L^\infty$ and $\lambda_1, \lambda_2 \in \mathbb{R}$,
2. for every $f \in L^\infty$, we obtain $\|Uf\|_{L^\infty} \leq \|f\|_{L^\infty}$,
3. for $f \in L^1$ and $g \in L^\infty$, we obtain:

$$\langle Pf, g \rangle = \langle f, Ug \rangle.$$

Next we will give a demonstration of the properties of the Kooman operator:

1. Trivially for $f_1, f_2 \in L^\infty$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. We have
$U(\lambda_1 f_1 + \lambda_1 f_2) = (\lambda_1 f_1 + \lambda_1 f_2)(S(x)) = \lambda_1 f_1(S(x)) + \lambda_1 f_2(S(x)) = \lambda_1 U(f_1) + \lambda_1 U(f_2)$.

2. Follows immediately from the Definition of the norm. Since $|f(x)| \leq \|f\|_{L^\infty}$ a.e. implies that
$$|f(S)| \leq \|f\|_{L^\infty}$$
the latter inequality gives the equation. Since by $U(f(x)) = f(S(x))$.

3. We first check it with $g = 1_A$, then the left-hand side, becomes
$$\langle Pf, g\rangle = \int_{\mathcal{X}} Pf 1_A d\mu = \int_A Pf d\mu.$$
The right-hand side becomes
$$\langle f, Ug\rangle = \int_{\mathcal{X}} fU1_A d\mu = \int_{\mathcal{X}} f1_A(S) d\mu = \int_{S^{-1}(A)} f(x) d\mu$$
thus item 3 is equivalent to
$$\int_A Pf d\mu = \int_{S^{-1}(A)} f d\mu$$
by the linearity we conclude property 3.

3 STUDYING CHAOS WITH DENSITIES

In this chapter we will study three types of transformations and how they can be characterized through the type of convergence of the sequence $\{P^n f\}$ where$P$ is the operator of Frobenius-Perron with respect to the transformation and f is a probability density.

3.1 INVARIANT MEASURE AND MEASURE-PRESERVING TRANSFORMATIONS

We begin this section with the Definition that a transformation S measure preserving. The dynamic behavior of measure preserving transformations that is the theme of ergodic theory.

Definition 22 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a measurable transformation, then S is said to be measure preserving if:*

$$\mu(S^{-1}(A)) = \mu(A) \text{ for all } A \in \mathcal{A}.$$

We will alternately say that the measure μ is *invariant* under S if S is measure preserving. We observe that every transformation that preserves the measure is also a non-singular transformation.

In the previous chapter we saw that the Frobenius-Perron operator associated with a transformation gives us the evolution of the probability density within the dynamics of the dynamical system, but there are cases in which this density does not change over time, but rather it is keeps fixed, invariant. The following theorem tells us that this happens if and only if the dynamical system is a measure preserving transformation.

Theorem 23 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation, and P the Frobenius-Perron operator associated with S. Consider a* *non-negative $f \in L^1$, then a measure μ_f given by:*

$$\mu_f(A) = \int_A f(x) d\mu$$

is invariant if and only if f is a fixed point of P.

Proof Fist we show the "only if" portion. Assume μ_f is invariant, then by the Definition of an invariant measure:

$$\mu_f(A) = \mu_f(S^{-1}(A)) \text{ for all } A \in \mathcal{A}$$

or

$$\int_A f(x) d\mu = \int_{S^{-1}(A)} f(x) d\mu \text{ for all } A \in \mathcal{A}.$$

However, by the Definition of the Frobenius-Perron operator, we obtain

$$\int_{S^{-1}(A)} f(x) d\mu = \int_A P f d\mu,$$

comparing, we immediately have:

$$\int_A P f(x) d\mu = \int_{S^{-1}(A)} f(x) d\mu = \int_A f(x) d\mu,$$

then

$$\int_A P f(x) d\mu = \int_A f(x) d\mu \text{ for all} A \Rightarrow P f(x) = f(x).$$

Conversely, if $Pf(x) = f(x)$ for some $f \in L^1$ and $f \geq 0$, then from Definition of Frobenius-Perron operator, we obtain:

$$\int_{S^{-1}(A)} f(x) d\mu = \int_A P f(x) d\mu = \int_A f(x) d\mu,$$

then $\mu_f(S^{-1}(A)) = \mu_f(A)$.

■

Remark: Note that the original measure μ is invariant if and only if $P1 = 1$.

There is a very important result, regarding the invariant measures, if we have a dynamical system $S : \mathcal{X} \to \mathcal{X}$, where S is a measurable function and we have that the μ measure is finite and invariant with respect to S, then we can guarantee that at some point almost every points that departed from a set with arbitrary positive measure, at some point large enough they will return to the set from which they left, this result is known as Poncaré recurrence theorem, see in [6] Theorem 1.2.1 page 4.

Now we will give some examples of measures that are invariant and other measures that are not invariant.

Example 24 *Consider the r-adic transformation the example 16*

$$S(x) = rx \mod 1$$

where we calculate the Frobenius-Perron operator associated to S is

$$Pf(x) = \frac{1}{r}\sum_{i=0}^{r-1} f\left(\frac{1}{r} + \frac{x}{r}\right)$$

thus

$$P1 = \frac{1}{r}\sum_{i=0}^{r-1} 1 = \frac{1}{r}(\underbrace{1 + \ldots + 1}_{r}) = 1$$

and by our previous, the Borel measure is invariant under the r-adic transformation.

Example 25 *Again consider the measure space $([0,1], \mathcal{B}, \mu)$ where μ is the Borel measure. Let $S : [0,1] \to [0,1]$ defined by $S(x) = 4x(1-x)$ in example 15 we saw that your Frobenius-Perron associate operator is*

$$Pf(x) = \frac{1}{4\sqrt{1-x}}\left[f\left(\frac{1}{2} - \frac{1}{2}\sqrt{1-x}\right) + f\left(\frac{1}{2} + \frac{1}{2}\sqrt{1-x}\right)\right]$$

then

$$P1 = \frac{1}{2\sqrt{1-x}}$$

so that the Borel measure μ is not invariant under S ,to find the invariant measure, we have to calculate a function that is Frobenius-Perron's fixed point operator or *equivalently solve*

$$f(x) = \frac{1}{4\sqrt{1-x}}\left[f\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-x}\right)+f\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-x}\right)\right]$$

This problem was first solved by Ulam and von Neumann [1947] who showed that the solution is given by

$$f_*(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

so by Theorem 23 the invariant measure by S corresponds to

$$\mu_*(A) = \int_A \frac{dx}{\pi\sqrt{x(1-x)}}$$

Example 26 *Now let $\mathcal{X}$ be the unit square in a plane, which we denote by $\mathcal{X} = [0,1]\times[0,1]$. The $\sigma-$algebra $\mathcal{B}$ is now generated by all possible rectangles of the form $[0,a]\times[0,b]$ and the Borel measure μ is the unique measure on $\mathcal{B}$ such that*

$$\mu([0,a]\times[0,b]) = ab.$$

We define the Baker transformation $S : \mathcal{X} \to \mathcal{X}$ by

$$S(x,y) = \begin{cases} (2x, \frac{1}{2}y) & \textit{if } 0 \geq x < \frac{1}{2} \textit{ and } 0 \leq y \leq 1 \\ (2x-1, \frac{1}{2}y+\frac{1}{2}) & \textit{if} \frac{1}{2} \leq x \leq 1 \textit{ and } 0 \leq y \leq 1 \end{cases}$$

Now we calculate the Frobenius-Perron operator for the Baker transformation.

$$S^{-1}([0,x]\times[0,y]) = \left[0, \frac{1}{2}x\right]\times[0,2y]$$

so from equation

$$Pf(x,y) = \frac{\partial^2}{\partial x \partial y}\int_0^{\frac{x}{2}}\int_0^{2y} f(s,t)dsdt = f\left(\frac{1}{2}x, 2y\right), \textit{ with } 0 \leq y < \frac{1}{2}$$

In the second case, form $\frac{1}{2} \leq y \leq 1$, we find that

$$S^{-1}([0,x]\times[0,y]) = \left(\left[0,\frac{1}{2}x\right]\right)\cup\left(\left[\frac{1}{2},\frac{1}{2}+\frac{1}{2}\right]\times[0,2y-1]\right)$$

hence

$$\begin{aligned} Pf(x,y) &= \frac{\partial^2}{\partial x \partial y}\left[\int_0^{\frac{x}{2}}\int_0^1 f(s,t)dsdt + \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{x}{2}}\int_0^{2y-1} f(s,t)dsdt\right] \\ &= f\left(\frac{1}{2}+\frac{1}{2}x, 2y-1\right) \end{aligned}$$

with $\frac{1}{2} \leq y \leq 1$. *Thus, finally*

$$Pf(x,y) = \begin{cases} f\left(\frac{1}{2}x, 2y\right) & \text{if } 0 \leq y < \frac{1}{2} \\ f\left(\frac{1}{2} + \frac{1}{2}x, 2y-1\right) & \text{if } \frac{1}{2} \leq y \leq 1 \end{cases}$$

so that $P1 = 1$, *and the Borel measure is, therefore, invariant under the Baker transformation.*

Example 27 *The Baker transformation is the previous example may be considered to be prototype of very important class of transformations originally introduced by Anosov [1963]. On of the simplest of the the Anosov diffeomorphims is given by*

$$S(x,y) = (x+y, x+2y) \mod 1$$

$$S^{-1}(x,y) = (2x-y, y-x) \mod 1$$

and thus

$$Pf(x,y) = f(2x-1, y-x)$$

is clear that $P1 = 1$, *which corresponds to the fact that* S *preserves the Borel measure.*

3.2 ERGODIC, MIXING, AND EXACT TRANSFORMATIONS

In this section we will study three types of dynamical systems, the ergodic, mixining and exact and their relationship with the Frobenius-Perron operator.

We start studying ergodic systems. The word ergodic comes from the Greek *ergos* which means work and *edos* which in turn means path, was introduced by the Austrian physicist Ludwig Boltzmann in his work on the kinetic theory of gases. Bolzmann thought that *the orbits cover the entire energy hyper surface constant*, that is, there is only one orbit. Later they baptized this hypothesis as an ergodic hypothesis, later in 1913 Michel Plancherel and Artur Rosenthal in [5] they demonstrate the impossibility of ergodic hypothesis based on the argument that a continuous curve that does not self-cross cannot fill a hypersurface of dimension greater than one. This led physicists to formulate a weaker condition at that Boltzman had formulated, which they called a quasi-ergodic hypothesis which was *the orbits are distributed densely on the hyper surface of constant energy. So it's not like there is only one orbit, but any orbit has as adhesion to all the hi-surface energy constant* All these concerns called the attention of the mathematicians Poincaré, Birkhoff and Neumann, who began to formulate and prove a series of theorems that we now know as Birkhoff's ergodic theorem and Neumann's ergodic theorem, ergodic theory began to take its first steps.

Next we will give the Definition of an ergodic system.

Definition 28 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation, then S is called ergodic if every invariant set $A \in \mathcal{A}$ is either $\mu(A) = 0$ or $\mu(\mathcal{X} \setminus A) = 0$.*

We can interpret ergodic systems, as systems where there is no subset with positive measure that are still during the evolution of the system.

The following result is an equivalence of the ergodic systems, which will be used later in the demonstration of Theorem 39.

Theorem 29 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation. S is ergodic if and only if for every measure function $f : \mathcal{X} \to \mathbb{R}$:*

$$f(S(x)) = f(x)$$

for almost all $x \in \mathcal{X}$, implies that f is constant almost everywhere.

Proof: We first show that ergodicty implies that f is constant. we obtain a function f satisfying $f(S(x)) = f(x)$ for almost all $x \in \mathcal{X}$, which is not constant almost everywhere, and that S is ergodic. Then there is some $r \in \mathcal{R}$ such that the sets

$A = \{x \in \mathcal{X}; f(x) \leq r\}$ and $B = \{x \in \mathcal{X}; f(x) > r\}$, have positive measure. These sets are also invariant because:

$$\begin{aligned} S^{-1}(A) &= \{x \in \mathcal{X}; S(x) \in A\} \\ &= \{x \in \mathcal{X}; f(S(x)) \leq r\} \\ &= \{x \in \mathcal{X}; f(x) \leq r\} = A \end{aligned}$$

and similarly for B, then S is not ergodic, which is a contradiction, thus, every f satisfying $f(S(x)) = f(x)$ for almost all $x \in \mathcal{X}$, must be constant.

To prove the converse, assume that S is not ergodic, then by Definition 28, there is a nontrivial set $A \in \mathcal{A}$ that is invariant. Let $f(x) = 1_A$, and since A is nontrivial, f is not a constant function. Moreover, since $A = S^{-1}(A)$ we obtain:

$$f(S(x)) = 1_A(S(x)) = 1_{S^{-1}(A)}(x) = 1_A = f(x) \quad a.e$$

and $f(S(x)) = f(x)$ for almost all $x \in \mathcal{X}$ is satisfied by a non constant function.

■

We may reformulate the previous theorem in term of the Koopman operator, as

Corollary 29.1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular, and U the Koopman operator with respect yo S, then S is ergodic if and only if all fixed points of U are constant functions.*

Theorem 30 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a non-singular transformation, and P the Frobenius-Perron operator associated with S. If S is ergodic, then there is at most one stationary density f_* of P. Further, if there is a unique stationary density f_* of P and $f_*(x) > 0$ a.e then S is ergodic.*

Proof: To prove the first part of Theorem assume that S is ergodic and f_1 and f_2 are different stationary densities of P. Let $g = f_1 - f_2$ so:

$$Pg = P(f_1 - f_2) = Pf_1 - Pf_2 = f_1 - f_2 = g$$

by Proposition 4

$$Pg^+ = g^+ \text{ and } Pg- = g^-.$$

Since by assumption, f_1 and f_2 are not only different but are also densities we obtain

$$g^+ \neq 0 \text{ and } g^- \neq 0$$

Let $A = Supp(g^+) = \{x \in \mathcal{X}; g^+(x) > 0\}$ and $B = Supp(g^-) = \{x \in \mathcal{X}; g^-(x) > 0\}$ it is evident that A and B are disjoint set and both have positive measure. By $g^+ \neq 0$ and $g^- \neq 0$ and Proposition 17 we obtain

$$A \subset S^{-1}(A) \text{ and } B \subset S^{-1}(B)$$

The sets $S^{-1}(A)$ and $S^{-1}(A)$ are disjoint, since:

$$A \cap B = \emptyset \Rightarrow S^{-1}(A) \cap S^{-1}(B) = \emptyset.$$

By induction, therefore, we obtain

$$A \subset S^{-1}(A) \subset S^{-2}(A) \subset \cdots \subset S^{-n}(A)$$

and

$$B \subset S^{-1}(B) \subset S^{-2}(B) \subset \cdots \subset S^{-n}(B)$$

Where $S^{-1}(A) \cap S^{-1}(B)$ are also disjoint for all n. Now define two set by

$$\widehat{A} = \textstyle\bigcup_{n=0}^{\infty} S^{-n}(A) \text{ and } \widehat{B} = \textstyle\bigcup_{n=0}^{\infty} S^{-n}(B)$$

these two sets $\widehat{A}$ and $\widehat{B}$ also disjoint and, furthermore they are invariant, because

$$S^{-1}(\widehat{A}) = S^{-1}\left(\bigcup_{n=0}^{\infty} S^{-n}(A)\right) = \bigcup_{n=1}^{\infty} S^{-n}(A) = \bigcup_{n=0}^{\infty} S^{-n}(A) = \widehat{A},$$

analogy with $\widehat{B}$. Neither $\widehat{A}$ and $\widehat{B}$ are of measure zero since A and B are not measure zero. Thus $\widehat{A}$ and $\widehat{B}$ are nontrivial invariant sets, which contradicts the ergodicty of S. Thus first portion of Theorem is proved.

To prove the second portion of Theorem, assume that $f_* > 0$ is the unique density satisfying $Pf_* = f_*$, but that S is not ergodic. If S is not ergodic then there exist a nontrivial set A such that $S^{-1}(A) = A$ and with $B = \mathcal{X} \setminus A$ where $S^{-1}(B) = B$, with these two set A and B we may white

$$f_* = 1_A f_* + 1_B f_*$$

so that

$$1_A f_* + 1_B f_* = P(1_A f_*) + P(1_B f_*).$$

The function $1_B f_*$ is equal to zero in the set $\mathcal{X} \setminus B = A = S^{-1}(A)$. Thus by Proposition 18 $P(1_B f_*) = 0$ in $A = \mathcal{X} \setminus B$, and likewise $P(1_A f_*) = 0$ in $B = \mathcal{X} \setminus A$, thus $1_A f_* + 1_B f_* = P(1_A f_*) + P(1_B f_*)$ implies that $1_A f_* = P(1_A f_*)$ and $1_B f_* = P(1_B f_*)$. Since f_* is positive on A and B, we may replace $1_A f_*$ by

$$f_A = \frac{1_A f_*}{\|1_A f_*\|}$$

then

$$P(f_A) = P\left(\frac{1_A f_*}{\|1_A f_*\|}\right) = \frac{1}{\|1_A f_*\|} P(1_A f_*) = \frac{1_A f_*}{\|1_A f_*\|} = f_A.$$

and $1_B f_*$ by

$$f_B = \frac{1_B f_*}{\|1_B f_*\|}$$

then

$$P(f_B) = P\left(\frac{1_B f_*}{\|1_B f_*\|}\right) = \frac{1}{\|1_B f_*\|} P(1_B f_*) = \frac{1_B f_*}{\|1_B f_*\|} = f_B.$$

this implies that there exist two stationary densities of P, which is in contradiction to our assumption. Thus, if there is a unique positive stationary density f_* of P, then S is ergodic. The following result is known as Birkhoff's ergodic theorem. This result can be obtained as a particular case of the subadditive ergodic theorem, see Theorem 3.3.3 of [6] page 78.

Theorem 31 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ a measurable transformation, and $f : \mathcal{X} \to \mathbb{R}$ an integrable function, if the measure μ is invariant, then there exist an integrable function f^* such that:*

$$f^*(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(S^k(x))$$

for almost all $x \in \mathcal{X}$. Moreover, the function f^* defined in this way is integrable and satisfies

$$\int f^*(x)d\mu = \int f(x)d\mu.$$

With the notion of ergodicity we may derive an important and often quoted extension of the Birkhoff pointwise ergodic theorem.

Theorem 32 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space, $S : \mathcal{X} \to \mathcal{X}$ be ergodic transformation. Then for any intergrable f, the average of f along the trajectory of S is equal almost everywhere to average of f over the space $\mathcal{X}$, that is:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(S^k(x)) = \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} f(x)d\mu.$$

Proof From $f^*(x) = f^*(S(x))$ for almost all $x \in \mathcal{X}$ and theorem 29 it follows that f^* is constant almost everywhere. Hence, from, when $\mu(\mathcal{X}) < \infty$:

$$\int_{\mathcal{X}} f^*(x)d\mu = \int_{\mathcal{X}} f(x)d\mu.$$

we obtain:

$$\int_{\mathcal{X}} f^*(x)d\mu = f^* \int_{\mathcal{X}} d\mu = f^*\mu(\mathcal{X}) = \int_{\mathcal{X}} fd\mu,$$

so that:

$$f^*(x) = \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} f(x)d\mu \quad a.e$$

■

Now we will study mixing systems, as in the case of ergodic systems, this concept also comes from physics. Before giving the Definition, let's see an intuitive idea of these systems.

The following example is based on the introductory example on page 34 [3] page 19

Suppose we are at a party after having had an ergodic theory test, and for fun we want to mix coke with some rum. While we look at the glass we see how the alcohol of the rum is mixed with the coke and we ask ourselves, given a portion B of coke in this glass, which is the weight of rum that will have the portion B, while we are mixing everything with a spoon? consider that this phenomenon is modeled a by discrete dynamical system S : glass$\rightarrow$glass, and also suppose that all the rum starts from an initial portion (initial set) that we will call A, then considering a measure of probability μ, we have the percentage of rum particles in B, it is given by

$$\frac{\mu(A \cap S^{-n}(B))}{\mu(B)}$$

where $A \cap S^{-n}(B)$ corresponds to the starting points $x \in A$ such that $S^n(x) \in B$. As expected within a time large enough, that the rum is mixed evenly throughout the coke, that is, if we had an initial percentage of 20% rum in relation to the whole mixture and we drink a little of this mixture, we will have that 20% of that mixture corresponds to rum. This means that

$$\frac{\mu(A \cap S^{-n}(B))}{\mu(B)} \rightarrow \mu(A) \text{ as } n \rightarrow \infty$$

or equivalent

$$\lim_{n\rightarrow\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B).$$

that is to say, that the measure of the points that start from B and reach A, tends to the product of the measure of A and B. This last result corresponds to the Definition of a mixing system. Now we will give the formal Definition.

Definition 33 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a normalized measure space and $S : \mathcal{X} \rightarrow \mathcal{X}$ a measure-preserving transformation. S is called mixing if*

$$\lim_{n\rightarrow\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B)$$

for all $A, B \in \mathcal{A}$.

The mixing dynamic systems, are actually a special type of ergodic dynamical systems. This fact can be demonstrated from the Corollary 3.3.2 of [7] page 48. However, this fact can also be observed later with Theorem 39.

Now we are going to study another type of dynamical system. Imagine that now we want to study, the probability of finding a particle of a red gas inside a sealed container, to study its evolution we will take a series of successive photos in regular times. We will assume that in principle the gas starts to expand from a small region A within this container, for the initial state, the probability of finding a gas particle outside region A is zero, but as time goes by passing the gas is expanding throughout the container, and for a time that is sufficiently large, we will have the probability of finding a particle anywhere is 1, taking this idea as a base we will define the exact dynamical systems, let's consider the Definition given by Rochlin in *Exact endomorphisms of Lebesgue spaces, Amer. Math. Soc. Transl. 2 (1964), 1-36.*

Definition 34 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a normalized measure space and $S : \mathcal{X} \to \mathcal{X}$ a measure-preserving transformation such that $S(A) \in \mathcal{A}$ for each $A \in \mathcal{A}$. if*

$$\lim_{n\to\infty} \mu(S^n(A)) = 1$$

for every $A \in \mathcal{A}$ and $\mu(A) > 0$, then S is called exact.

A first observation that we can make about the exact systems, is that, this system cannot be invertible, in fact, if we assume otherwise, we would have the following, for any invetible measure preserving transformation S, we have $\mu(S(A)) = \mu(S^{-1}(S(A))) = \mu(A)$ and by induction $\mu(S^n(A)) = \mu(A)$, which violate Definition 34.

Proposition 35 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ let a normalized measure space and let $S : \mathcal{X} \to \mathcal{X}$ be measure preserving. Then S is exact if and only if*

$$\mathcal{A}^T = \bigcap_{n=0}^{\infty} S^{-n}(\mathcal{A})$$

Consist of the set of measure 0 or 1.

Proof Let us assume that $A \in \mathcal{A}^T$, with $0 < \mu(A) < 1$ and let $A_n \in \mathcal{A}$ be such that $A = S^{-n}(A_n)$ for $n = 1, 2, \ldots$. Since S preserve μ, we have $\mu(A_n) = \mu(A)$ for $n = 1, 2, \ldots$. We also have $S^n(A) = S(S^{-n}A_n) \subset A_n$. Hence, $\mu(S^n(A)) \leq \mu(A) < 1$ for $n = 1, 2, \ldots$, which contradicts the exactness of S.
Let $A \in \mathcal{A}$ and $\mu(A) > 0$. If $\lim_{n\to\infty} \mu(S^n A) < 1$, we may assume that for some $a < 1$, we have

$$\mu(S^n A) \leq a < 1, \text{ for } n = 1, 2, \ldots.$$

For any $n \geq 0$ we have $S^{-n}(S^n A) \subset S^{-(n+1)}(S^{n+1}A)$. Thus, the set $B = \bigcup_{n=0}^{\infty} S^{-n}(S^n(A))$ belongs to $\mathcal{A}^T$. Since $A \subset B$, and $< 0\mu(A) \leq \mu(B) = 1$. On the other hand,

$$\mu(B) = \lim_{n\to\infty} \mu(S^{-n}(S^n A) = \lim_{n\to\infty} \mu(S^n(A)) \leq a < 1$$

■

Then we will give the Definition of three forms of convergence, which will be used in Theorem 39.

Definition 36 *A sequence of functions $\{f_n\}$ with $f_n \in \mathcal{L}^p$, $1 \leq p < \infty$ is (weakly) cesàro convergent to $f \in \mathcal{L}^p$ if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \langle f_k; g \rangle = \langle f; g \rangle \textit{ for all } g \in \mathcal{L}^{p^*}.$$

Definition 37 *A sequence of functions $\{f_n\}$ with $f_n \in \mathcal{L}^p$, $1 \leq p < \infty$ is weakly convergent to $f \in \mathcal{L}^p$ if*

$$\lim_{n\to\infty} \langle f_n; g \rangle = \langle f; g \rangle \textit{ for all } g \in \mathcal{L}^{p^*}.$$

Definition 38 *A sequence of functions $\{f_n\}$ with $f_n \in \mathcal{L}^p$, $1 \leq p < \infty$ is strongly convergent to $f \in \mathcal{L}^p$ if*

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

The following theorem characterizes the type of transformation $S : \mathcal{X} \to \mathcal{X}$ (*ergodic, mixing or exact*) through the convergence mode (*Cesàro, weak or strong*) of the sequence $\{P^n f\}$ for any density function f, where P is the operator of Frobenius-Perron associated with the transformation S.

Theorem 39 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a normalized measure space and $S : \mathcal{X} \to \mathcal{X}$ a measure-preserving transformation, and P the Frobenius-Perron operator corresponding to S. then*

1. *S is ergodic if and only if the sequence $\{P^n f\}$ is Cesàro convergent to 1 for all $f \in \mathcal{D}$,*
2. *S is mixing if and only if $\{P^n f\}$ is weakly convergent to 1 for all $f \in \mathcal{D}$,*
3. *S is exact if and only if $\{P^n f\}$ is strongly convergent to 1 for all $f \in \mathcal{D}$.*

Proof

1. Let S is ergodic, then by Theorem 32 we have:

$$\frac{1}{n}\sum_{k=0}^{n-1}\langle P^k f, g\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle f, U^k g\rangle = \langle f, \frac{1}{n}\sum_{k=0}^{n-1} U^k g\rangle$$

take $n \to \infty$ we obtain:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\langle P^k f, g\rangle = \langle f, \int_{\mathcal{X}} g d\mu\rangle = \langle f, 1\rangle\langle 1, g\rangle = \langle 1, g\rangle.$$

We're going to take a fixed point from the Koopman operator and we're going to show that this point is a constant function, which will finally prove that S is ergodic by Corollary 29.1. Let g an fixed point of the Koopman operator, then:

$$\frac{1}{n}\sum_{k=0}^{n-1}\langle P^n f, g\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle f, U^n g\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle f, g\rangle = \langle f, g\rangle.$$

Then we obtain the following equation:

$$\langle f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle, \quad \forall f \in \mathcal{D}, g \in \mathcal{L}^{\infty}.$$

let's show now, that the last equation above implies that S is ergodic. Rewriting the equation, we obtain:

$$\int_{\mathcal{X}} fgd\mu = \int_{\mathcal{X}} gd\mu, \quad \forall f \in \mathcal{D}, \text{ and } g \in \mathcal{L}^{\infty}$$

Let $f = \frac{\varphi}{\int_{\mathcal{X}} \varphi d\mu}$ with $\int_{\mathcal{X}} \varphi d\mu \neq 0$, then

$$\int_{\mathcal{X}} \varphi gd\mu = \int_{\mathcal{X}} \varphi d\mu \int_{\mathcal{X}} gd\mu, \quad \forall \varphi \in \mathcal{L}^1 \text{ with } \int_{\mathcal{X}} \varphi d\mu \neq 0 \text{ and } , g \in \mathcal{L}^{\infty}$$

Let $\varphi = 1_A$, then:

$$\int_A gd\mu = \mu(A) \int_{\mathcal{X}} gd\mu, \quad g \in \mathcal{L}^{\infty}$$

Consider the following set $A = \left\{ x \in \mathcal{X}; g(x) = \int_{\mathcal{X}} gd\mu \right\}$. We suppose that $\mu(A) < 1$, then

$$\mu(A) \int_A gd\mu < \int_A gd\mu = \mu(A) \int_{\mathcal{X}} gd\mu = \mu(A) g(x)$$

then

$$\mu(A) \int_A gd\mu < \mu(A) g(x) \Rightarrow \int_A gd\mu < g(x)$$

That is a contradiction to the Definition of our set A, then $\mu(A) = 1$ and $g(x) = \int_{\mathcal{X}} gd\mu$ q.t.p., therefore g is constant q.t.p. of this last result we conclude by lemma 29.1 that S is ergodic.

2. Assume S is mixing, which by Definition, means

$$\lim_{n \to \infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B)$$

for all $A, B \in \mathcal{A}$, we can be rewritten in integral form as

$$\lim_{n \to \infty} \int_{\mathcal{X}} 1_A 1_B(S^n(x)) d\mu = \int_{\mathcal{X}} 1_A d\mu \int_{\mathcal{X}} 1_B d\mu.$$

By applying the Definition of the Koopman operator and the scalar product to this equation, we obtain

$$\lim_{n \to \infty} \langle 1_A, U^n 1_B \rangle = \langle 1_A, 1 \rangle \langle 1, 1_B \rangle.$$

Since the Koopman operator is adjoin to Frobenius-Perron operator the equation above may be rewritten as

$$\lim_{n \to \infty} \langle P^n 1_A, 1_B \rangle = \langle 1_A, 1 \rangle \langle 1, 1_B \rangle.$$

or

$$\lim_{n\to\infty}\langle P^n f, 1_B\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

for $f = 1_A$ and $g = 1_B$. Since this relation holds for characteristic function it must also hold for the simple functions

$$f = \sum_i \lambda_i 1_{A_i} \text{ and } g = \sum_i \sigma 1_{b_i}.$$

Further, every function $g \in \mathcal{L}^\infty$ is the uniform limit of simple functions $g_k \in \mathcal{L}^\infty$, and every function $f \in \mathcal{L}^1$ is the strong (in $\mathcal{L}^1$ norm) limit of a sequence of simple function $f_k \in \mathcal{L}^1$. Obviously,

$$\begin{aligned}|\langle P^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle| \leq |\langle P^n f, g\rangle - \langle P^n f_k, g_k\rangle| + |\langle P^n f_k, g_k\rangle - \langle f_k, 1\rangle\langle 1, g_k\rangle| \\ + |\langle f_k, 1\rangle\langle 1, g_k\rangle - \langle f, 1\rangle\langle 1, g\rangle|\end{aligned}$$

If $\|f_k - f\| \leq \varepsilon$ and $\|g_k - g\|_{\mathcal{L}^\infty} \leq \varepsilon$, then the fist and last terms on the right-hand side the equation above satisfy:

$$\begin{aligned}|\langle P^n f, g\rangle - \langle P^n f_k, g_k\rangle| &\leq |\langle P^n f, g\rangle - \langle P^n f_k, g\rangle| + |\langle P^n f_k, g\rangle - \langle P^n f_k, g_k\rangle| \\ &= |\langle P^n f - P^n f_k, g\rangle| + |\langle P^n f_k, g - g_k\rangle| \\ &= \left|\int_{\mathcal{X}} (P^n f - P^n f_k) g d\mu\right| + \left|\int_{\mathcal{X}} P^n f_k (g - g_k) d\mu\right| \\ &= \int_{\mathcal{X}} |P^n f - P^n f_k||g| d\mu + \int_{\mathcal{X}} |P^n f_k|g - g_k| d\mu \\ &\leq \| P^n f - P^n f_k \| \| g \|_{\mathcal{L}^\infty} + \| P^n f_k \| \| g - g_k \|_{\mathcal{L}^\infty} \\ &\leq \varepsilon \| g \|_{\mathcal{L}^\infty} + \varepsilon \|f_k\| \leq \varepsilon(\| g \|_{\mathcal{L}^\infty} + \|f\| + \varepsilon)\end{aligned}$$

and analogously

$$|\langle f_k, 1\rangle\langle 1, g_k\rangle - \langle f, 1\rangle\langle 1, g\rangle| \leq \varepsilon(\| g \|_{\mathcal{L}^\infty} + \|f\| + \varepsilon).$$

Thus these terms are arbitrary small ε. Finally, for fixed k the middle term

$$|\langle P^n f_k, g_k\rangle - \langle f_k, 1\rangle\langle 1, g_k\rangle|$$

converges to zero as $n \to \infty$ Remember that f_k, g_k are simple functions, which shows that the right-hand side of inequality can be as small as we wish for large

n. This completes the proof that mixing implies the convergence of $\langle P^n, g\rangle$ to $\langle f, 1\rangle\langle 1, g\rangle = \langle 1, g\rangle$ for all $f \in \mathcal{D}$ and $g \in \mathcal{L}^\infty$. Conversely, this convergence implies the mixing condition. As for all $f \in \mathcal{D}$ and $g \in \mathcal{L}^\infty$ we have

$$\{P^n f\} \xrightarrow{w} 1 \text{ if and only if } \lim_{n\to\infty} \langle P^n f, g\rangle = \langle 1, g\rangle$$

Set $f = \dfrac{1_A}{\mu(A)}$ and $g = 1_B$, then

$$\lim_{n\to\infty} \langle P^n \frac{1_A}{\mu(A)}, 1_B\rangle = \mu(A) \lim_{n\to\infty} \langle P^n 1_A, 1_B\rangle$$

then

$$\lim_{n\to\infty} \langle P^n 1_A, 1_B\rangle = \mu(A)\langle 1, g\rangle = \mu(A)\mu(B).$$

On the other hand $\langle P^n 1_A, 1_B\rangle = \langle 1_A, U^n 1_B\rangle = \mu(A \cap S^n(B))$. Hence

$$\lim_{n\to\infty} \mu(A \cap S^n(B)) = \mu(A)\mu(B)$$

that is to say that S is a mixing.

3. Lastly, Assume S is exact. The $\sigma-$algebras $S^{-n}(\mathcal{A})$ form a decreasing sequence of $\sigma-$algebras. Since S is exact, the $\sigma-$algebra $\mathcal{A}^T = \bigcap_{n=1}^{\infty} S^{-n}(\mathcal{A})$ consist of sets of measure 0 or 1. By Proposition and Proposition

$$P^n f \circ S^n = E(f|S^{-n}(\mathcal{A})) \to E(f|S^{-n}(\mathcal{A}^T))$$

in $\mathcal{L}^1$ as $n \to \infty$. Since $\mathcal{A}^T$ consists of sets of measure 0 or 1, then $E(f|S^{-n}(\mathcal{B}^T)) = \int f d\mu = 1$. Thus, we have

$$\int_{\mathcal{X}} |P^n f \circ S^n - 1| d\mu \to 0 \text{ as } n \to \infty.$$

But

$$\int_{\mathcal{X}} |P^n f \circ S^n - 1| d\mu = \int_{\mathcal{X}} |P^n f \circ S^n - 1 \circ S^n| d\mu = \int_{\mathcal{X}} |P^n f - 1| d\mu$$

Thus $P^n f \to 1$ in $\mathcal{L}^1$ as $n \to \infty$.

We show that strong convergence of $P^n f \xrightarrow{s} 1$ for all $f \in \mathcal{D}$ implies exactness, let's start with $\mu(S^n(A))$ with $\mu(A) > 0$ as the space is normalized, then $0 \leq \mu(S^n(A)) \leq 1$, Let $f = \dfrac{1_A}{\mu(A)} \in \mathcal{D}$ then

$$\begin{aligned}\mu(S^n(A)) = \int_{S^n(A)} 1 d\mu &\geq \left|\int_{S^n(A)} P^n f d\mu\right| - \left|\int_{S^n(A)} P^n f - 1 d\mu\right| \\ &\geq \left|\int_{S^n(A)} P^n f d\mu\right| - \|P^n f - 1\| \\ &\geq 1 - \|P^n f - 1\|\end{aligned}$$

Since $\int_{S^n(A)} P^n f d\mu = \int_{S^{-n}(S^n(A))} P^n f d\mu = \int_A f d\mu = \dfrac{1}{\mu(A)} \int_A 1_A d\mu = \dfrac{\mu(A)}{\mu(A)} =$ 1. Hence take $n \to \infty$ we have $\mu(S^n(A)) \to 1$.

$\blacksquare$

Note that, since P is linear, convergence of $\{P^n f\}$ to 1 for $f \in \mathcal{D}$ is equivalent to the convergence of $\{P^n f\}$ to $\langle , f\rangle$ for every $f \in \mathcal{L}^1$. This observation is, of course, valid for all types of convergence: Cesaro, weak, and strong. Thus we may restate Theorem 39 in the equivalent form.

Theorem 40 *Under the assumptions of Theorem 39, the following equivalence hold:*

1. *S is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \langle P^k f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle, \text{ for } f \in \mathcal{L}^1, g \in \mathcal{L}^\infty,$$

2. *S is mixing if and only if*

$$\lim_{n\to\infty} \langle P^k f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle, \text{ for } f \in \mathcal{L}^1, g \in \mathcal{L}^\infty,$$

3. *S is exact if and only if*

$$\lim_{n\to\infty} \|P^k f - \langle f, 1\rangle\| = 0, \text{ for } f \in \mathcal{L}^1.$$

Next, we will give three examples of dynamical systems and use Theorem 39 to say if they are ergodic, mixing or exact.

Example 41 *One of the first examples given during the course of ergodic theory corresponds to the rotation of the unit circle, there are many ways to write this function of rotation with angle ϕ, but in particular we will define it as follows. Let $S : [0, 2\pi) \to [0, 2\pi)$ defined by*

$$S(x) = x + \phi \mod 2\pi$$

when $\frac{\phi}{2\pi}$ is rational, we have that S is not ergodic, because if we consider a rotation angle for example $\phi = 3$ and consider a small arc of the unit cycle, which obviously has a positive measure, when we iterate S three times, we will get the same Initial set and passing we have obtained two sets with positive measure, let's call A_1, A_2 and A_3 those sets. Now consider a new set that will be the union of the three previous sets, as we can see when we apply S on this new set, we will have all the points of A_1 pass to A_2, all the points of A_2 pass to A_3 and all points of A_3 pass to A_1, which means that after this process we have obtained the same set, that is, that the set formed by the union of these three sets, is invariant and not only does that also have positive measure, and in a way analogous to any ϕ such that $\frac{\phi}{2\pi}$ is rational. From this we conclude that when $\frac{\phi}{2\pi}$ is rational S is not ergodic.

Here we proved that it is ergodic when $\frac{\phi}{2\pi}$ is irrational. It is straightforward to show that S preserves the Borel measure μ and the normalized measure $\frac{\mu}{2\pi}$. We take as our linearly dense set in $\mathcal{L}^p([0, 2\pi])$, that consisting of the functions $\{sin(kx), cos(lx); k, l = 0, 1, \ldots\}$. We will show that, for each function g belonging to this set, GIO

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} U^k g(x) = 1$$

uniformly for all x, thus implying by theorem 39. To simplify the calculations, note that

$$\sin kx = \frac{e^{ikx} - e^{-ikx}}{2i}, \cos kx = \frac{e^{ikx} + e^{-ikx}}{2}$$

where $i = \sqrt{-1}$. Consequently, it is sufficient to verify only for $g(x) = \exp(ikx)$ with k an arbitrary (not necessary positive) integer.
We have, for $k \neq 0$,

$$U^l g(x) = g(S^l(x)) = e^{ik(x+l\phi)},$$

so that

$$u_n(X) = \frac{1}{n}\sum_{l=0}^{n-1} U^l g(x),$$

obeys

$$\begin{aligned} u_n(x) &= \frac{1}{n}\sum_{l=0}^{n-1} e^{ik(x+l\phi)} \\ &= \frac{1}{n}e^{ikx}\frac{e^{ink\phi}-1}{e^{ik\phi}-1} \end{aligned}$$

and

$$\begin{aligned} \|u_n\|_{\mathcal{L}^2} &\leq \frac{1}{n(e^{ik\phi}-1)}\left(\int_0^{2\pi} |e^{ikx}(e^{ink\phi}-1)|^2 \frac{dx}{2\pi}\right)^{\frac{1}{2}} \\ &\leq \frac{2}{n(e^{ik\phi}-1)}. \end{aligned}$$

Thus $u_n(x)$ converges in $\mathcal{L}^2$ to zero. Also, however, with our choice of $g(x)$,

$$\langle 1, g\rangle = \int_0^{2\pi} e^{ikx}\frac{dx}{2\pi} = \frac{1}{ik}(e^{2\pi ik}-1) = 0$$

and condition of Theorem 39 for ergodicity is satisfied with $k \neq 0$. When $k = 0$ the calculation is even simpler, sin $g(x) = 1$ and thus $u_n = 1$. Noting also that

$$\langle 1, g\rangle = \int_0^{2\pi} \frac{dx}{2\pi} = 1$$

we have again that $u_n(x)$ converge to $\langle 1, g\rangle$.

Example 42 *We demonstrate the exactness of the $r-$adic transformation, that we saw in the example 16*

$$S(x) = rx \mod 1$$

It sufficient to demonstrate that $\{P^n f\}$ converge strongly to 1 with $f \in \mathcal{D}$, then

$$Pf(x) = \frac{1}{r}\sum_{i=0}^{r-1} f\left(\frac{i}{r}+\frac{x}{r}\right).$$

and thus by induction

$$P^n f(x) = \frac{1}{r^n}\sum_{i=0}^{r^n-1} f\left(\frac{i}{r^n}+\frac{x}{r^n}\right).$$

However, in the limit as $n \to \infty$, the right-hand side of this equation approaches the Riemann integral of f over $[0,1]$, that is

$$\lim_{n\to\infty} P^n f(x) = \int_0^1 f(s)ds = 1, \text{ uniformly in } x.$$

Example 43 *Here we show that the Anosov diffeomorphism*

$$S(x,y) = (x+y, x+2y) \mod 1$$

is mixing, it is sufficient to show that $U^n g(x,y) = g(S^n(x,y))$ converges weakly to 1, for g in linearly dense set in $\mathcal{L}^p([0,1]\times[0,1])$.
Observe that for $g(x,y)$ periodic in x and y with period 1, $g(S(x,y)) = g(x+y, x+2y), g(S^2(x,y)) = g(2x+3y, 3x+5y)$, and so on. By induction we easily find that

$$U^n g(x,y) = g(a_{2n-2}x + a_{2n-1}y, a_{2n-1}x + a_{2n}y)$$

where the a_n are the Fibonacci numbers given by $a_0 = a_1 = 1$ $a_{n+1} = a_n + a_{n-1}$. Thus, if take $g(x,y) = \exp[2\pi i(kx+ly)]$ and $f(x,y) = \exp[-2\pi i(px+qy)]$, then we have

$$\langle f, U^n g\rangle = \int_0^1\int_0^1 \exp[2\pi i(ka_{2n-1} + la_{2n-1} - p)x + (ka_{2n-1} + la_{2n} - q)]dsdy$$

and it is straightforward to show that

$$\langle f, U^g\rangle = \begin{cases} 1 & \text{if } (ka_{2n-2} + la_{2n-1} - p) = (ka_{2n-1} + la_{2n} - q) = 0 \\ 0 & \text{if otherwise} \end{cases}$$

Now we show that for large n either

$$ka_{2n-2} + la_{2n-1} - p \text{ or } ka_{2n-1} + la_{2n} - q$$

is different from zero if at least one of k,l,p,q is different from zero. If $k = l = 0$ but $p \neq 0$ or $q \neq 0$ this obvious. We may suppose that either k or l is not zero. Assume $k \neq 0$ and that $ka_{2n-2} + la_{2n-1} - p = 0$ for infinitely many n. Thus

$$k\frac{a_{2n-2}}{a_{2n-1}} + l - \frac{p}{a_{2n-1}} = 0$$

It is well known that

$$\lim_{n\to\infty} \frac{a_{2n-2}}{a_{2n-1}} = \frac{2}{1+\sqrt{5}}$$

and

$$\lim_{n\to\infty} a_n = \infty$$

hence

$$\lim_{n\to\infty} \left(k\frac{a_{2n-2}}{a_{2n-1}} + l - \frac{p}{a_{2n-1}} \right) = \frac{2k}{1+\sqrt{5}} + l.$$

However, this limit can never be zero because k *and* l *are integers. Thus* $ka_{2n-2} + la_{2n-1} - p \neq 0$ *for large* n*. Therefore, for large* n*,*

$$\langle f, U^n \rangle = \begin{cases} 1 & \text{if } k = p = q = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \langle 1, g \rangle &= \int_0^1 \int_0^1 \exp[2\pi(kx + ly)]dxdy \\ &= \begin{cases} 1 & \text{if } k = l = 0 \\ 0 & \text{if } k \neq 0 \text{ or } l \neq 0 \end{cases} \end{aligned}$$

so that

$$\begin{aligned} \langle f, 1 \rangle \langle 1, g \rangle &= \int_0^1 \int_0^1 \langle 1, g \rangle \exp[-2\pi(px + qy)]dxdy \\ &= \begin{cases} \langle 1, g \rangle & \text{if } p = q = 0 \\ 0 & \text{if } p \neq 0 \text{ or } q \neq 0 \end{cases} \\ &= \begin{cases} 1 & \text{if } k = l = p = q = 0 \\ 0 & \text{if otherwise} \end{cases} \end{aligned}$$

Thus

$$\langle f, U^n g \rangle = \langle f, 1 \rangle \langle 1, g \rangle$$

for large n*, as a consequence,* $\{U^n g\}$ *converge weakly to* $\langle 1, g \rangle$*. Therefore, mixing of the Anosov diffeomorphism is demonstrated.*

4 THE ASYMPTOTIC PROPERTIES OF DENSITIES

In this chapter we are going to study sufficient conditions for the existence of stationary densities for Markov operators and we will also study a special type of markov operators that we call contrustives, and we will use the spectral decomposition theorem for constructive markov operators, to demonstrate that all Construction markov operators have a stationary density.

4.1 WEAK PRECOMPACTNESS

In this section we are going to introduce the weak precompacts or also called weak sequancially compacts and we will just mention some results, which will be occupied in the next sections of this chapter.

Definition 44 *The set $\mathcal{F}$ called weakly precompact if every sequence of functions $\{f_n\}$ with $f_n \in \mathcal{F}$, contains a weakly convergence subsequence $\{f_{a_n}\}$ that converges to an $f \in \mathcal{L}^p$.*

Proposition 45 *A set of function $\mathcal{F} \subset \mathcal{F}^1$, $\mu(\mathcal{X}) < \infty$, is weakly precompact if and only if:*

1. *There is an $M < \infty$ such that $\|f\| \leq M$ for all $f \in \mathcal{F}$, and*
2. *For every $\varepsilon > 0$ there is a $\delta > 0$ such that*

$$\int_A |f(x)| d\mu < \varepsilon, \text{ if } \mu(A) < \delta \text{ and } f \in \mathcal{F}.$$

Proof See Corollary 11 page 294 of Theorem 7 of the page 291 of [8].

Corollary 45.1 *Let $g \in \mathcal{L}^1$ be a nonnegative function and $\mu(\S) < \infty$. Then the set of all functions $f \in \mathcal{L}^1$ such that*

$$|f(x)| \leq g(x) \text{ for } x \in \mathcal{X} \text{ a.e}$$

is weakly precompact in $\mathcal{L}^1$.

Corollary 45.2 *Let $M > 0$ be a positive number and $p > 1$ be given. If $\mu(\S) < \infty$, then the set of all function $f \in \mathcal{L}^1$ such that*

$$\|f\|_{\mathcal{L}^p} \leq M$$

is weakly precompact in $\mathcal{L}^1$.

4.2 PROPERTIES OF THE AVERAGES $A_N F$

We will start this part by defining the average function and assume, for simplicity, that we are always working on a measure space $(\mathcal{X}, \mathcal{A}, \mu)$, whether $P : \mathcal{L}^1 \to \mathcal{L}^1$ is a Markov operator, then defined the Cesàro average of P by

$$A_n f = \frac{1}{n} \sum_{k=0}^{n-1} P^k f$$

We will show that every Markov operator whose average sequence is weakly precompact, has a fixed point, we will see later that this is simply a consequence of the Hans-Banach theorem. To demonstrate this, let's start with one with some Propositions.

Proposition 46 *Let $P : \mathcal{L}^1 \to \mathcal{L}^1$ an Markov operator, then for all $f \in \mathcal{L}^1$, we have*

$$\lim_{n \to \infty} \|A_n f - A_n P f\| = 0$$

Proof Let's start the demonstration by first calculating for an $f \in \mathbb{L}^1$ fix

$$A_n f - A_n P f = \frac{1}{n} \sum_{k=0}^{n-1} P^k f - \frac{1}{n} \sum_{k=0}^{n-1} P^{k+1} f = \frac{1}{n}(f - P^n f)$$

and thus

$$\|A_n f - A_n P f\| \leq \frac{1}{n} \left(\|f\| + \|P^n\|\right) \leq \frac{2}{n} \|f\|$$

since P is a Markov operator and we saw in Proposition 2 of chapter 1 that for an arbitrary function in $\mathcal{L}^1$ we have that $\|P^n f\| \leq \|f\|$ taking $n \to \infty$ in the inequality above, we have

$$\|A_n f - A_n f\| \leq \frac{2}{n} \|f\| \to 0$$

■

Proposition 47 *Let P an linear operator in $\mathcal{L}^1$, if, for $f \in \mathcal{L}^1$, there is a subsequence $\{A_{a_n}\}$ of sequence $\{A_n\}$ that converges weakly to $f_* \in \mathcal{L}^1$, then $Pf_* = f_*$.*

Proof Note that by linearity of P, we have

$$PA_{a_n}f = P\left(\frac{1}{a_n}\sum_{k=0}^{a_n-1} P^k f\right) = \frac{1}{a_n}\sum_{k=0}^{a_n-1} P^k Pf = A_{a_n}Pf$$

thus $PA_{a_n}f = A_{a_n}Pf$, from here, we have that $A_{a_n}Pf$ weakly converges to Pf_*. Since $\{A_{a_n}Pf\}$ has the same limit as $\{A_n f\}$, then $Pf_* = f_*$

■

The following result is a consequence of the second geometric form of the Hahn-Banach Theorem

Proposition 48 *Let M a closed subspace of a normed space E. Then, for all $x_0 \in E \setminus M$ exist a functional $\varphi \in E^*$, such that $\varphi(x_0) = 1$ and $\varphi(x) = 0$ for all $x \in M$.*

Proof See Corollary 3.4.10 page 74 of [1].

Lemma 49 *Let $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator. If $\{A_n f\}$ is weakly precompact, then for every $\varepsilon > 0$, the function $f - f_*$ can be written as*

$$f - f_* = Pg - g + r.$$

Where $g \in \mathcal{L}^1$ and $\|r\| < \varepsilon$.

Proof Suppose that for some ε there does not exist an r such that it suffers the equation of the lemma. then we have

$$f - f_* \notin \overline{(P - I)\mathcal{L}^1(\mathcal{X})}$$

The central idea of this demonstration, is to use one of the consequences, of the Hanh-Banach theorem that we enunciate above, for this we must first prove that $\overline{(P-I)\mathcal{L}^1(\mathcal{X})}$ is a linear space.

Let's prove that $\overline{(P-I)\mathcal{L}^1(\mathcal{X})}$ is a linear space. Indeed, Let $u, v \in \overline{(P-I)\mathcal{L}^1(\mathcal{X})}$, then exist two sequence such that

$$u = \lim_{n\to\infty}(PU_n - U_n), \text{ with } U_n \in \mathcal{L}^1$$
$$v = \lim_{n\to\infty}(PV_n - V_n), \text{ with } V_n \in \mathcal{L}^1.$$

Let $\lambda \in \mathbb{R}$ Then

$$\begin{aligned}\lambda u + v =& \lambda \lim_{n\to\infty}(PU_n - U_n) + \lim_{n\to\infty}(PV_n - V_n)\\ =& \lim_{n\to\infty}(P\lambda U_n - \lambda U_n + PV_n - V_n)\\ =& \lim_{n\to\infty}(P(\lambda U_n + V_n) - (\lambda U_n + V_n)) \in \overline{(P-I)\mathcal{L}^1(\mathcal{X})}.\end{aligned}$$

Since $\lambda U_n + V_n \in \mathcal{L}^1$. Then by 48, there must exist a $g_0 \in \mathcal{L}^\infty$ such that

$$\langle f - f_*, g_0\rangle \neq 0$$

and $\langle h, g_0\rangle = 0$ for all $h \in \overline{(P-I)\mathcal{L}^1(\mathcal{X})}$. In particular

$$\langle (P-I)P^j f, g_0\rangle = 0.$$

thus

$$\langle P^j f, g_0\rangle = \langle P^j f, g_0\rangle \text{ for } j = 0, 1, \ldots$$

and by induction we must, therefore, have

$$\langle P^j f, g_0\rangle = \langle f, g_0\rangle.$$

As a consequence

$$\frac{1}{n}\sum_{j=0}^{n-1}\langle P^j f, g_0\rangle = \frac{1}{n}\sum_{j=0}^{n-1}\langle f, g_0\rangle = \langle f, g_0\rangle$$

or

$$\langle A_n f, g_0\rangle = \langle f, g_0\rangle.$$

Let $\{A_{a_n}\}$ a subsequece of $\{A_n\}$ such that converge weakly to f_*, we have

$$\lim_{n\to\infty} \langle A_{a_n} f, g_0\rangle = \langle f_*, g_0\rangle$$

joining this result with the one above we have

$$\langle f, g_0\rangle = \langle f_*, g_0\rangle$$

which gives

$$\langle f - f_*, g_0\rangle = 0$$

what is a contradiction, then, then for every $\varepsilon > 0$, the function $f - f_*$ can be written as

$$f - f_* = Pg - g + r.$$

Where $g \in \mathcal{L}^1$ and $\|r\| < \varepsilon$.

Theorem 50 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator. If for a given $f \in \mathcal{L}^1$, the sequence $\{A_n f\}$ is weakly precompact, then it converges strongly to some $f_* \in \mathcal{L}^1$ that is a fixed point of P, namely $Pf_* = f_*$. Furthermore, if $f \in \mathcal{D}$, then $f_* \in \mathcal{D}$, so that f_* is stationary density.*

Proof Because $\{A_n f\}$ is weakly precompact by assumption, there exists a subsequence $\{A_{a_n}\}$ that converges weakly to some $f_* \in \mathcal{L}^1$. Further, by Proposition 47, we know $Pf_* = f_*$. Write $f \in \mathcal{L}^1$ in the form

$$f = (f - f_*) + f_*$$

for lemma 49 we have, for every $\varepsilon > 0$ exist $g \in \mathcal{L}^1$ and $\|r\| < \varepsilon$ such that, the function $f - f_*$ can be written as

$$f - f_* = Pg - g + r.$$

or equivalent

$$f = Pg - g + r + f_*$$

Applying the average in the expression above

$$A_n f = A_n(Pg - g) + A_n r + A_n f_*$$

Then estimating the following expression through triangular inequality

$$\|A_n f - f_*\| = \|A_n(f - f_*)\| \leq \|A_n(Pg - g)\| + \|A_n r\|.$$

By Proposition 46 we know that $\|A_n(Pg - g)\|$ is strongly convergent to zero as $n \to \infty$ and by our assumption $\|A_n r\| \leq \|r\| < \varepsilon$

$$\|A_n r\| = \|\frac{1}{n}\sum_{k=0}^{n-1} P^k r\| \leq \frac{1}{n}\sum_{k=0}^{n-1} \|P^k r\| \leq \frac{1}{n}\sum_{k=0}^{n-1} \|r\| = \|r\|.$$

Thus, for sufficiently large n, we must have

$$\|A_n f - f_*\| \leq \varepsilon$$

Since ε is arbitrary, this proves that $\{A_n f\}$ is strongly convergence to f_*. To show that if $f \in \mathcal{D}$, then $f_* \in \mathcal{D}$. By Definition the densities we have $f \geq 0$ and $\|f\| = 1$, then by Definition the Markov operator we have $Pf \geq 0$ and $\|Pf\| = 1$, so that $P^n f \geq 0$ and $\|P^n f\| = 1$. As consequence $A_n f \geq 0$ and

$$\|A_n f\| = \int_{\mathcal{X}} |A_n f| d\mu = \int_{\mathcal{X}} A_n f d\mu = \frac{1}{n}\sum_{k=0}^{n-1} \int_{\mathcal{X}} P^n f d\mu = 1$$

And, since $\{A_n f\}$ is strongly convergent to f_*, we must have $f_* \in \mathcal{D}$. This complete the proof.

■

Corollary 50.1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator, If, for some $f \in \mathcal{D}$ there is a $g \in \mathcal{L}^1$ such that*

$$P^n f \leq g$$

for all n, then there is an $f_* \in \mathcal{D}$ such that $Pf_* = f_*$, that is, f_* is a stationary density.

Proof By assumption $P^n f \le g$ so that

$$0 \le A_n f = \frac{1}{n}\sum_{k=0}^{n-1} P^K f \le g$$

and, thus, $|A_n f| \le g$. By applying our first criterion for weak precompactness, we know that $\{A_n f\}$ is weakly precompact. Then Theorem 50 completes the argument.

■

Corollary 50.2 *Again let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator, If, for some $f \in \mathcal{D}$ there is $M > 0$ and $p > 1$ such that*

$$\|P^n f\|_{\mathcal{L}^p} \le M$$

for all n, then there is an $f_* \in \mathcal{D}$ such that $Pf_* = f_*$.

Proof We have

$$\|A_n f\|_{\mathcal{L}^p} = \|\frac{1}{n}\sum_{k=0}^{n-1} P^k f\|_{\mathcal{L}^p} \le \frac{1}{n}\sum_{k=0}^{n-1} \|P^k f\|_{\mathcal{L}^p} \le \frac{1}{n}(nM) = M.$$

Hence, by our second criterion for weak precompactness, $\{A_n f\}$ is weakly precompact, and again Theorem 50 completes the proof.

■

Theorem 51 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator with a unique stationary f_*. If $f_*(x) > 0$ for all $x \in \mathcal{X}$, then*

$$\lim_{n\to\infty} A_n f = f_* \text{ for all } f \in \mathcal{D}$$

Proof First assume $\dfrac{f}{f_*}$ is bounded. By setting $c = \sup\left(\dfrac{f}{f_*}\right)$, we have

$$P^n f \le P^n c f_* = cP^n f_* = cf_*$$

and

$$A_n f \le cA_n f_* = cf_*$$

thus the sequence $\{A_n f\}$ is weakly precompact and, by Theorem 50, is convergent to a stationary density. Since f_* is the unique stationary density, then $\{A_n f\}$ must converge to f_*. Thus Theorem is proved when $\dfrac{f}{f_*}$ is bounded.

In the general case, write $f_c = \min(f, cf_*)$, We then have

$$f = \frac{1}{\|f_c\|} f_c + r_c$$

where

$$r_c = \left(1 - \frac{1}{\|f_c\|}\right) f_c + f - f_c.$$

Since $f_*(x) > 0$, exist any $c \in \mathbb{R}$ for each x such that

$$f(x) \leq cf_*(x)$$

thus $\lim\limits_{n\to\infty} f_c(x) = f(x)$. By Definition of f_c, we have $f_c \leq f$, thus by the Lebesgue dominated convergence theorem, $\|f_c - f\| \to 0$ and $\|f_c\| \to \|f\| = 1$ as $c \to \infty$. Thus we have that

$$\|r_c\| = \|\left(1 - \frac{1}{\|f_c\|}\right) f_c + f - f_c\| = \|f - \frac{f_c}{\|f_c\|}\| \to 0 \text{ as } c \to \infty$$

that is to say that $\|r_c\|$ converges strongly to zero as $c \to \infty$. By choosing $\varepsilon > 0$ we can find a value c such that $\|r_c\| < \dfrac{\varepsilon}{2}$, so for this fixed value of c, we have

$$\|A_n r_c\| = \|\frac{1}{n}\sum_{k=0}^{n-1} P^k r_c\| \leq \|r_c\| < \frac{\varepsilon}{2}$$

However, since $\frac{f_c}{\|f_c\|}$ is a density bounded by $\frac{cf_c}{\|f_c\|}$. according to the first part of the proof,

$$\|A_n\left(\frac{1}{\|f_c\|} f_c\right) - f_*\| \leq \frac{\varepsilon}{2}$$

for sufficiently large n. Obtain

$$\|A_n f - f_*\| \leq \varepsilon$$

for sufficiently large n.

Corollary 51.1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a normalized measure space, $S : \mathcal{X} \to \mathcal{X}$ a measure preserving transformation, and P the corresponding Frobenius-Perron operator. Then S is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f = 1 \text{ for every } f \in \mathcal{D}$$

Proof The proof is immediate. Since S is measure preserving, we have $P1 = 1$. If S is ergodic, then by Theorem , $f_*(x) = 1$ is the unique stationary density of P and, by Theorem 51, the convergence of follows. Conversely, it the convergence of holds, applying to stationary density f gives $f = 1$. Thus $f_*(x) = 1$ is the unique stationary density of P and again, by Theorem , the transformation S is ergodic

■

4.3 ASYMTOTIC PERIODICITY OF $\{P^N F\}$

Definition 52 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a finite measure space. A Markov operator P is called constructive if there exist a $\delta > 0$ and $\kappa < 1$ such that for every $f \in \mathcal{D}$ there is an integer $n_0(f)$ for which*

$$\int_E P^n d\mu \leq \kappa \text{ for } n \geq n_0(f) \text{ and } \mu(E) \leq \delta$$

Note that for every density f the integral in inequality is bounded above by one. Thus condition for constrictiveness means the eventually $(n \leq n_0(f))$ this integral cannot be close to one for sufficiently small set E. This clearly indicates that constrictiveness rules out the possibility that P^f is eventually concentrated on a set very small or vanishing measure.

In [4] in the remark 5.1.3, mentions an interesting equivalence, relating to compact set. *A markov operator $P : \mathcal{L}^1 \to \mathcal{L}^1$ is contrictive if and only if there exists a compact set $\mathcal{F} \subset \mathcal{L}^1$ such that*

$$\lim_{n\to\infty} dist(P^n f, \mathcal{F}) = 0$$

for any $f \in \mathcal{D}$.

If the space $\mathcal{X}$ is not finite, we wish to have a Definition of constrictivess that also prevents $P^n f$ from begin dispersed throughout the entire space. To accomplish this we extend Definition

Definition 53 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be a* σ*-finite measure space. A Markov operator* P *is called constrictive if there exists* $\delta > 0$*, and* $\kappa < 1$ *and a measurable set* B *of finite measure, such that for every density* f *there is an integer* $n_0(f)$ *for which*

$$\int_{(\mathcal{X}\setminus B)\cup E} P^n f d\mu \leq \kappa \text{ for } n \leq n_0(f) \text{ and } \mu(E) \leq \delta.$$

From the Definition, one might think that verifying constrictiveness is difficult since it is required to find two constants δ and κ as well as a set B with rather specific properties. Howevwe, it is often rather easy ti verify constrictiveness using one the two following Propositions.

Proposition 54 *Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be a finite measure space and* $P : \mathcal{L}^1 \to \mathcal{L}^1$ *be a Markov operator. Assume there is a* $P > 1$ *and* $K > 0$ *such that for every density* $f \in \mathcal{D}$ *we have* $P^n f \in \mathcal{L}^p$ *for sufficiently large* n*, and*

$$\limsup_{n\to\infty} \|P^n f\|_{\mathcal{L}^p} \leq K.$$

Then P *is constrictive.*

Proof From $\limsup_{n\to\infty} \|P^n f\|_{\mathcal{L}^p} \leq K$, there is an integer $n_0(f)$ such that

$$\|P^n f\|_{\mathcal{L}^p} \leq K + 1 \text{ for } n \geq n_0(f).$$

Thus, by criteria 2 the family $\{P^n f\}$, for $n \geq n_0(f)$, $f \in \mathcal{D}$, is weakly precompact. Finally, for fixed $\varepsilon \in (0, 1)$, criteria 3, implies there is $\delta > 0$ such that

$$\int_E P^n f(x) d\mu < \varepsilon \text{ if } \mu(E) < \delta$$

Thus P is constrictiveness.

■

Our next Proposition may be even more useful in demonstrated the contrictiveness of an operator P

Proposition 55 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a $\sigma-$finite measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ be a Markov operator. If there exist an $h \in \mathcal{L}^1$ and $\lambda < 1$ such that*

$$\limsup_{n\to\infty} \|(P^n f - h)^+\| \leq \lambda \text{ for } f \in \mathcal{D},$$

then P is constrictive.

Proof Let $\varepsilon = \frac{1}{4}(1-\lambda)$ and take $\mathcal{F} = \{h\}$. Since $\mathcal{F}$, which contains only one element, in evidently weakly precompact (it is also strongly precompact, but this property is no useful to us here). then bt criterion 3, there exist a $\delta > 0$ such that

$$\int_E h(x)d\mu < \varepsilon, \text{ for } \mu(E) < \delta.$$

Furthermore, by there is a measurable set B of finite measure for which

$$\int_{\mathcal{X}\setminus B} h(x)d\mu < \varepsilon.$$

Now fix $f \in \mathcal{D}$. From we may choose an integer $n_0(f)$ such that

$$\|(P^n - h)^+\| \leq \lambda + \varepsilon \text{ for } n \geq n_0(f),$$

and, as a consequence,

$$\int_C P^n f(x)d\mu \leq \int_C h(x)d\mu + \lambda + \varepsilon$$

for an arbitrary set $C \in \mathcal{A}$. Setting $C = (\mathcal{X} \setminus B) \cup E$ in and using and we have

$$\int_{\mathcal{X}\setminus B)\cup E} P^n f(x)d\mu \leq \int_{\mathcal{X}\setminus B} + \int_E h(x)d\mu + \lambda + \varepsilon < 3\varepsilon + \lambda = 1 - \varepsilon$$

this completes the proof.

■

Theorem 56 (spectral decomposition theorem) *Let P be a constrictive Markov operator. Then there is an integer r, two sequence of nonnegative functions $g_i \in \mathcal{D}$ and $\kappa_i \in \mathcal{L}^\infty$, $i = 1, \ldots, r$ and an operator $Q : \mathcal{L}^1 \to \mathcal{L}^1$ such that for every $f \in \mathcal{L}^1$, Pf may be written in the form*

$$Pf(x) = \sum_{i=1}^{r} \lambda_i(f) g_1(x) + Qf(x) \tag{4.1}$$

where

$$\lambda_i(f) = \int_{\mathcal{X}} f(x) k_i(x) d\mu.$$

The functions g_i and operator Q have the following properties

1. *$g_i(x)g_j(x) = 0$ for all $i \neq j$, so that functions g_i have disjoint supports.*
2. *For each integer i there exist an unique integer $\alpha(i)$ such that $Pg_i = g_{\alpha(i)}$. Further $\alpha(i) \neq \alpha(j)$ for $i \neq j$ and thus operator P just serves to permute the functions g_i.*
3. *$\|P^n Q f\| \to 0$ as $n \to \infty$ for every $f \in \mathcal{L}^1$.*

Proof See page 141 of [9].

■

Remark From representation 4.1 of Theorem 56 for Pf, it immediately follows that the structure of $P^{n+1}f$ is given by

$$P^{n+1}f(x) = \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^n(i)}(x) + Q_n f(x) \tag{4.2}$$

where $Q_n = P^n Q$ and $\alpha^n(i) = \alpha(\alpha^{n-1}) = \ldots$, and $\|Q_n f\| \to 0$ as $n \to \infty$. The terms under the summation in 4.2 are just permuted with each application of P, and since r is finite the sequence

$$\sum_{i=1}^{r} \lambda_i(f) g_{\alpha^n(i)}(x)$$

must be periodic with a period $\tau \leq r!$. Since $\{\alpha^n(1), \ldots, \alpha^n(r)\}$ is simply a permutation of $\{1, \ldots, r\}$, there is a unique i corresponding to each $\alpha^n(i)$. Thus it is clear that the summation above may be be rewritten as

$$\sum_{i=1}^{r} \lambda_{\alpha^{-n}(i)}(f) g_i(x),$$

where $\{\alpha^{-n}(i)\}$ denote the inverse permutation $\{\alpha^n(i)\}$.

Rewritten the summation in this form clarifies how sucessive applications od operator P really work. Since the funcions g_i are supported on disjoint sets, each successive application of operator P leads to a a new set of scaling coefficients $\lambda_{\alpha^{-n}})(f)$ associated with each functions $g_i(x)$.

A sequence $\{P^n\}$ for witch formula 4.1 is satisfied will be called *asymptotically periodic*. Using this notion. Theorem 56 may be rephrased as follows: If P is a constrictive operator, then $\{P^n\}$ is asymptotically periodic.

It is actually rather easy to obtain an upper bound on the integer r appearing in equation 4.1 if we can find an upper bound function for $P^n f$. Assume that P is a Markov operator and there exists a function $h \in \mathcal{L}^1$ such that

$$\lim_{n \to \infty} \|(P^n f - h)^+\| = 0 \text{ for } f \in \mathcal{D}$$

Thus P is constrictive and representation 4.1 is valid. Let τ be the period of sequence 5.1 , so that, from 4.1 and, we have

$$Lf(x) = \lim_{n \to \infty} P^{nr} = \sum_{i=1}^{r} \lambda_i(f) g_i(x) \leq h(x) \text{ for } f \in \mathcal{D}.$$

Set $f = g_k$ so that $Lf(x) = g_k(x) \leq h(x)$. By integrating over the support of g_k, bearing in mind that the supports of the g_k are disjoint, and summing from $k = 10$ to $k = r$, we have

$$\sum_{k=1}^{r} \int_{Supp_{g_k}} g_k(x) d\mu \leq \sum_{k=1}^{r} \int_{Supp_{g_k}} h(x) d\mu \leq \|h\|.$$

Since $g_k \in \mathcal{D}$, this reduces to

$$r \leq \|h\|$$

which is the desired result.

Now we will use the result above to show that every construction Markov operator admits a stationary density

Proposition 57 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a constrictive Markov operator. Then P has a stationary density.*

Proof Let a density f be defined by

$$f(x) = \frac{1}{r} \sum_{i=1}^{r} g_i(x),$$

where r and g_i were defined in Theorem 56. Because of property (2) Theorem 56,

$$Pf(x) = \frac{1}{r} \sum i = 1^r g_{\alpha(i)}(x)$$

and thus $Pf = f$, which completes the proof.

■

5 ENTROPY

The concept of entropy was first introduced by Clausius and later used in a different form by Boltzmann in his pioneering work on the kinetic theory of gases published in 1866. since then, entropy has played a pivotal role in the development of many areas in physics and chemistry and has had important ramification in ergodic theory.

In this chapter we will work with the Bolzamann Gibbs entropy defined in the context of statistical mechanics, we will study their behavior in relation to sequence $\{P^n f\}$, where P is a Markov operators and later when P is an Operators of Frobenius-Perron for $f \in \mathcal{D}$ and finally we will study how Bolzamann Gibbs entropy behavior can say if a dynamical system is exact. Next we will give the Definition.

Definition 58 *Let$f \geq 0$ and $\eta(f) \in L^1$ where $\eta(u) = -u \log u$ and $\eta(0) = 0$, then the entropy of f is defined by*

$$\mathcal{H}(f(x)) = \int_{\mathcal{X}} \eta(f(x)) d\mu \tag{5.1}$$

If $\mu(\mathcal{X}) < \infty$, then the integral is always well defined for every $f \leq 0$. In fact, the integral over the positive parts of $\eta(f)$, $(\eta(f))^+ = \max(0, \eta(f))$ is always finite, thus. $\mathcal{H}(f)$ is either finite or equal to $-\infty$. Since we take $\eta(0) = 0$, the function $\eta(u)$ is continuous for all $u \geq 0$. Inequality that we will use much in this chapter is the Gibb's inequality

$$u - u\log(u) \leq v - u\log(v) \text{ for all } u, v \geq 0. \tag{5.2}$$

If f, g are two densities such that $\eta(f)$ and $f \log(g)$ are integrable, then from Gibb's inequality, we obtain

$$-\int_{\mathcal{X}} f(x) \log f(x) d\mu \leq -\int_{\mathcal{X}} f(x) \log g(x) d\mu \tag{5.3}$$

Proposition 59 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space with $\mu(\mathcal{X}) < \infty$, and consider all the possible densities f defined on $\mathcal{X}$, then, in the family of all such densities, the* *maximal entropy occurs for the constant density:*

$$f_0(x) = \frac{1}{\mu(\mathcal{X})} \tag{5.4}$$

Proof Pick an arbitrary $f \in \mathcal{D}$ so that entropy of f is given by

$$\mathcal{H}(f) = -\int_{\mathcal{X}} f(x)\log(f(x))d\mu$$

and inequality

$$\begin{aligned}\mathcal{H}(f(x)) &\leq -\int_{\mathcal{X}} f(x)\log(f_0(x))d\mu \\ &= -\log\left(\frac{1}{\mu(\mathcal{X})}\right)\int_{\mathcal{X}} f(x)d\mu\end{aligned}$$

or

$$\mathcal{H}(f(x)) \leq -\log\left(\frac{1}{\mu(\mathcal{X})}\right)$$

However, the entropy of f_0 is

$$\begin{aligned}\mathcal{H}(f_0) &= -\int_{\mathcal{X}} \frac{1}{\mu(\mathcal{X})}\log\left(\frac{1}{\mu(\mathcal{X})}\right) \\ &= -\log\left(\frac{1}{\mu(\mathcal{X})}\right).\end{aligned}$$

So $\mathcal{H}(f) \leq \mathcal{H}(f_0)$ for all $f \in \mathcal{D}$.

■

Corollary 59.1 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be measure space with $\mu(\mathcal{X}) < \infty$ and $f \in L^1$, then*

$$\mathcal{H}(f) \leq \log\mu(\mathcal{X}) \leq \frac{1}{e}\mu(\mathcal{X}).$$

Next we will give some examples of maximum entropy of a family of densities.

Example 60 *Let $X = [0, \infty)$ and consider all possibles densities f such that the fist moment of f is given by*

$$\int_0^\infty xf(x)dx = \frac{1}{\lambda} \tag{5.5}$$

then the density $f_0(x) = \lambda e^{-\lambda x}$ *maximizes the entropy.*

$$\begin{aligned}
\mathcal{H}(f) &\leq -\int_0^\infty f(x)\log(\lambda e^{-\lambda x})dx \\
&= -\log\lambda\int_0^\infty f(x)dx + \int_0^\infty \lambda x f(x)dx \\
&= -\log\lambda\underbrace{\int_0^\infty f(x)dx}_{1} + \underbrace{\lambda\frac{1}{\lambda}}_{1} \\
&= -\log\lambda + 1
\end{aligned}$$

and the other hand

$$\begin{aligned}
\mathcal{H}(f_0) &= -\int_0^\infty \lambda e^{-\lambda x}\log(\lambda e^{-\lambda x})dx \\
&= \left(e^{-\lambda x}(\log\lambda e^{-\lambda x} - 1)\right)_0^\infty \\
&= -\log\lambda + 1.
\end{aligned}$$

Example 61 *For our next example take* $X = (-\infty, \infty)$ *and consider all possible densities* $f \in \mathcal{D}$ *such that the second moment of* f *is finite, that is,*

$$\int_{-\infty}^{\infty} x^2 f(x)dx = \sigma^2 \tag{5.6}$$

then the maximal entropy is achieved for the Gaussian density

$$f_0(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}. \tag{5.7}$$

As before, we calculate that, for arbitrary $f \in \mathcal{D}$ *satisfying 5.6,*

$$\begin{aligned}
\mathcal{H}(f) &\leq -\int_{-\infty}^{\infty} f(x)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}\right)dx \\
&= -\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\int_{-\infty}^{\infty} f(x)dx + \frac{1}{2\sigma^2}\int_{-\infty}^{\infty} x^2 f(x)dx \\
&= \frac{1}{2} - \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)
\end{aligned}$$

Futher

$$\mathcal{H}(f_0) = \frac{1}{2} - \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)$$

so that the entropy is maximized with thw Gaussian

These two examples are simply special case covered by the following simple statement

Proposition 62 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space. Assume that a sequence $g_1, \ldots, g_m$ of measure function is given as well as two sequences of real constants $\overline{g}_1, \ldots, \overline{g}_m$ and $\nu_1, \ldots, \nu_m$ that satisfy*

$$\overline{g}_i = \frac{\int_{\mathcal{X}} g_i \exp[-\nu_i g_i] d\mu}{\int_{\mathcal{X}} \prod_{i=1}^{m} \exp[-\nu_i g_i] d\mu} \tag{5.8}$$

where all of the integrals are finite. Then the maximum of the entropy $\mathcal{H}(f)$ for all $f \in \mathcal{D}$, subject to the conditions

$$\overline{g}_i = \int_{\mathcal{X}} g_i f(x) d\mu, i = 1, \ldots, m \tag{5.9}$$

occurs for

$$f_0(x) = \frac{\prod_{i=1}^{m} \exp[-\nu_i g_i]}{\int_{\mathcal{X}} \prod_{i=1}^{m} \exp[-\nu_i g_i] d\mu} \tag{5.10}$$

Proof For simplicity, set

$$Z = \int_{\mathcal{X}} \prod_{i=1}^{m} \exp[-\nu_i g_i] d\mu$$

so

$$f_0(x) = Z^{-1} \prod_{i=1}^{m} \exp[-\nu_i g_i].$$

From inequality 5.3, we have

$$\begin{aligned}
\mathcal{H}(f) &\leq -\int_{\mathcal{X}} f(x) \log f_0(x) d\mu \\
&= -\int_{\mathcal{X}} f(x) \left(-\log Z - \sum_{i=1}^{m} \nu_i g_i \right) d\mu \\
&= \log Z + \sum_{i=1}^{m} \nu_i \int_{\mathcal{X}} f(x) g_i d\mu \\
&= \log Z + \sum_{i=1}^{m} \nu_i \overline{g}_i
\end{aligned}$$

Furthermore, it easy to show that

$$\mathcal{H}(f_0) = \log Z + \sum_{i=1}^{m} \nu_i \overline{g}_i$$

and thus $\mathcal{H}(f) \leq \mathcal{H}(f_0)$

■

Note that if $m = 1$ and $g(x)$ is identified as the energy of system, then the maximal entropy occurs for

$$f_0(x) = Z^{-1} e^{-\nu g(x)},$$

which is just the Gibbs canonical distribution function, with the partition function Z given by

$$Z = \int_{\mathcal{X}} e^{-\nu g(x)} d\mu.$$

Further, the maximal entropy

$$\mathcal{H}(f_0) = \log Z + \nu \overline{g}$$

is just the thermodynamic entropy. As is well known, all of the results of classical thermodynamic can derived with the partition function Z and the preceding entropy. Indeed, the contents of Proposition 62 have been extensively used by Jaynes and Katz in an alternative formulation and development of classical and quantum statistical mechanics.

5.1 ENTROPY OF $P^N F$ WHEN P IS A MARKOV OPERATOR

Theorem 63 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be finite measure space and $P : L^1 \to L^1$ a Markov operator. If P has a constant stationary density (P1=1), then:*

$$\mathcal{H}(Pf) \geq \mathcal{H}(f) \textit{ for all } f \geq 0 \textit{ and } f \in L^1.$$

Proof: integrating jensen's inequelity over entire spece $\mathcal{X}$.

$$\int_{\mathcal{X}} \eta(Pf)d\mu \geq \int_{\mathcal{X}} P\eta(f)d\mu = \int_{\mathcal{X}} \eta(f)d\mu$$

Since P preserve the integral. However, the left-most integral is $\mathcal{H}(Pf)$ and las integral is $\mathcal{H}(f)$. For a finite measure space, we know that the maximal entropy $H_{\max}$ is $-\log\left(\frac{1}{\mu(\mathcal{X})}\right)$, so that $-\log\left(\frac{1}{\mu(\mathcal{X})}\right) \geq \mathcal{H}(P^n f) \geq \mathcal{H}(f)$.

$\blacksquare$

This, in conjunction with theorem 63, tell us that in finite measure space when P has a constant stationart density, the entropy never decreases and is bounded above by $-\log\left(\frac{1}{\mu(\mathcal{X})}\right)$ and if $\mu(\mathcal{X}) = 1$ then $H_{\max} = 0$.

In the case of Markov operator without a constant stationary densities, it happen that the sequence $\mathcal{H}(P^n f)$ is not increasing as n increases.

5.2 ENTROPY $\mathcal{H}(P^N F)$ WHEN P IS A FROBENIUS-PERRON OPERATOR

Theorem 64 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be finite measure space and $S : \mathcal{X} \to \mathcal{X}$ be an invertible measure-preserving transformation. If P is the Frobenius-Perron operator corresponding to S. then*

$$\mathcal{H}(P^n f) = \mathcal{H}(f) \textit{ for all } n.$$

Proof If S is invertible and measure preserving, then by equation of lemma 18.1 we have $Pf(x) = f(S^{-1}(x))$ since $J^{-1} \equiv 1$. If P_1 is the Frobenius-Perron operator corresponding to S^{-1}, we also have $P_1 f(x) = f(S(x))$. Thus $P_1 P f = P P_1 f = f$ so $P_1 = P^{-1}$. From Theorem 63 we also have

$$\mathcal{H}(P_1 P f) \geq \mathcal{H}(Pf) \geq \mathcal{H}(f),$$

but, since $P_1 P f = P^{-1} P f = f$, we conclude that $\mathcal{H}(Pf) = \mathcal{H}(f)$, so $\mathcal{H}(P^n f) = \mathcal{H}(f)$ for all n

∎

Theorem 65 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be normalized measure space and $S : \mathcal{X} \to \mathcal{X}$ measure-preserving transformation and P is the Frobenius-Perron operator corresponding to S. If S is exact then*

$$\lim_{n\to\infty} \mathcal{H}(P^n f) = 0 \text{ for all } f \in \mathcal{D} \text{ such that } \mathcal{H}(f) > -\infty.$$

Proof Assume initially that f is bounded, that is, $0 \leq f \leq c$. Then

$$0 \leq P^n f \leq P^n c = cP^n 1 = c.$$

Without ant loss of generalty, we can assume that $c > 1$. Further, since $\eta(u) \leq 0$ for $u \geq 1$, we have

$$0 \leq \mathcal{H}(P^n f) \leq \int_{A_n} \eta(P^n(f(x))d\mu$$

where

$$A_n = \{x : 1 \leq P^n f(x) \leq c\}$$

Now, by the mean value theorem, we obtain

$$\begin{aligned}\left|\int_{A_n} \eta(P^n(f(x))d\mu\right| &= \int_{A_n} |\eta(P^n f(x) - \eta(1)|d\mu \\ &\leq k \int_{A_n} |P^n f(x) - 1|d\mu \\ &\leq k \int_{\mathcal{X}} |P^n f(x) - 1|d\mu = \|P^n - 1\|,\end{aligned}$$

where

$$k = \sup_{1\leq u\leq c} |\eta'(u)|.$$

Since S is exact, from Theorem, we have $\|P^n f - 1\| \to 0$ as $n \to \infty$ for all $f \in \mathcal{D}$ and thus

$$\lim_{n\to\infty} \int_{A_n} \eta(P^n(f(x))d\mu = 0$$

From inequality, it follows that $\mathcal{H}(P^n f)$ converges to zero. Now relax the assumption that f is bounded and write f in the form

$$f = f_1 + f_2$$

where

$$f_1(x) = \begin{cases} 0 & \text{if } f(x) > c \\ f(x) & \text{if } 0 \leq f(x) \leq c \end{cases}$$

and $f_2 = f - f_1$. Fixing $\varepsilon > 0$, we may choose c sufficiently large so that

$$\|f_2\| < \varepsilon \text{ and } \mathcal{H}(f_2) > -\varepsilon.$$

Write $P^n f$ in the form

$$P^n f = (1-\delta) P^n \left(\frac{1}{1-\delta} f_1 \right) + \delta P^n \left(\frac{1}{\delta} f_2 \right)$$

where $\delta = \|f_2\|$. Now $\dfrac{f_1}{1-\delta}$ is a bounded density, and so from the first part of our proof we know that for n sufficiently large

$$\mathcal{H}\left(P^n \left(\frac{1}{1-\delta} f_1 \right) \right) > -\varepsilon.$$

Furthermore

$$\begin{aligned} \delta \mathcal{H}\left(P^n \left(\frac{1}{\delta} f_2 \right) \right) &= \mathcal{H}(P^n f_2) - \log\left(\frac{1}{\delta} \right) \int_{\mathcal{X}} P^n f_2 d\mu \\ &= \mathcal{H}(P^n f_2) - \|f_2\| \log\left(\frac{1}{\delta} \right) \\ &= \mathcal{H}(P^n f_2) + \delta \log \delta. \end{aligned}$$

Since $\mathcal{H}(P^n f_2) \leq \mathcal{H}(f_2) > -\varepsilon$, this last expression becomes

$$\delta \mathcal{H}\left(P^n \left(\frac{1}{\delta} f_2 \right) \right) \geq -\varepsilon + \delta \log \delta.$$

Combining these results and inequality, we have

$$\begin{aligned} \mathcal{H}(P^n f) &\geq (1-\delta) \mathcal{H}\left(P^n \left(\frac{1}{1-\delta} f_1 \right) \right) + \delta \mathcal{H}\left(P^n \left(\frac{1}{\delta} f_2 \right) \right) \\ &\geq -\varepsilon(1-\delta) - \varepsilon + \delta \log \delta \end{aligned}$$

$$= -2\varepsilon + \delta\varepsilon + \delta \log \delta.$$

Since $\mu(\mathcal{X}) = 1$, we have $\mathcal{H}(P^n f) \leq 0$. Further since $\delta < \varepsilon$ and ε is arbitrary, the right-hand side of is also arbitrarily small, and Theorem is proved.

■

5.3 BEHAVIOR OF $P^N F$ FROM $\mathcal{H}(P^N F)$

Theorem 66 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space $\mu(X) < \infty$, and $P : \mathcal{L}^1 \to \mathcal{L}^1$ a Markov operator such that $P1 = 1$. If there exists a constant $c > 0$ such that for every bounded $f \in \mathcal{D}$*

$$\mathcal{H}(P^n f) \geq -c \text{ for } n \text{ sufficiently large,}$$

then P is constrictive

Proof Observe that $P1 = 1$ implies that Pf is bounded for bounded f. Thus, to proved our theorem, it is sufficient to show that the set $\mathcal{F}$ of all bounded $f \in \mathcal{D}$ that satisfy

$$\mathcal{H}(f) \geq -c$$

is weakly precompact. We will use criterion to demostrate the weak precompacteness of $\mathcal{F}$. Since $\|f\| = 1$ for all $f \in \mathcal{D}$, the fist part of the criterion is satisfied. To check the second part take $\varepsilon > 0$. Pick $l = e^{-1}\mu(\mathcal{X})$, $N = \exp \dfrac{2(c+l)}{\varepsilon}$ and $\delta = \frac{\varepsilon}{2N}$, and take a set $A \subset \mathcal{X}$ such that $\mu(A) < \delta$. Then

$$\int_A f(x)d\mu = \int_{A_1} f(x)d\mu + \int_{A_2} f(x)d\mu$$

where

$$A_1 = \{x \in A : f(x) \leq N\}$$

$$A_2 = \{x \in A : f(x) > N\}.$$

The fist integral on the right-hand side the above equation clearly satisfies

$$\int_{A_1} f(x)d\mu \leq N\delta = \frac{\varepsilon}{2}$$

In evaluating the second integral, note that from $\mathcal{H}(f) \geq -c$, it follows that

$$\begin{aligned}
\int_{A_2} f(x) \log f(x) d\mu &\leq c - \int_{\mathcal{X}\setminus\mathcal{A}_\in} f(x) \log f(x) d\mu \\
&\leq c + \int_{\mathcal{X}\setminus\mathcal{A}_\in} \eta_{\max} d\mu \\
&\leq c + \left(\frac{1}{e}\right) \mu(\mathcal{X}) = c + l.
\end{aligned}$$

Therefore

$$\int_{A_2} f(x) \log N d\mu < x + l$$

or

$$\int_{A_2} f(x) d\mu < \frac{c+l}{\log N} = \frac{\varepsilon}{2}$$

Thus

$$\int_A f(x) d\mu < \varepsilon$$

and $\mathcal{F}$ is weakly precompact. Thus, by Definition, the operator P is constrictive.

$\blacksquare$

Theorem 67 *Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space, $S : \mathcal{X} \to \mathcal{X}$ a measure-preserving transformation, and P the Frobenius-Perron operator corresponding to S. If*

$$\lim_{n\to\infty} \mathcal{H}(P^n f) = 0$$

for all bounded $f \in \mathcal{D}$, then S is exact.

Proof It follows from Theorem 66 that P is constructive. Furthermore since S is measure preserving, we know that P has a constant stationary density. From Proposition, we, therefore, have

$$P^{n+1} f(x) = \sum_{i=1}^{r} \lambda_{\alpha^{-n}(i)}(f) 1_{A_1}(x) + Q_n f(x) \text{ for } f \in \mathcal{L}^1$$

If we can demonstrate that $r = 1$, then form Theorem we will have shown S to be exact. Pick

$$f(x) = \left[\frac{1}{\mu(A_1)}\right] 1_{A_1}(x)$$

as an initial f. If τ is the asymptotic period of $P^n f$, then we must have

$$P^{n\tau} f(x) = \left[\frac{1}{\mu(A_1)}\right] 1_{A_1}(x).$$

However, by assumption

$$\lim_{n\to\infty} \mathcal{H}(P^n f) = 0$$

and, since the sequence $\{\mathcal{H}(P^{n\tau} f)\}$ is constant sequence, we must have

$$\mathcal{H}\left(\left[\frac{1}{\mu(A_1)}\right] 1_{A_1}\right) = 0.$$

Note that, by Proposition, $\mathcal{H}(f) = 0$ only if

$$f(x) = 1_{\mathcal{X}}(x)$$

So, clearly, we must have

$$\left[\frac{1}{\mu(A_1)}\right] 1_{A_1}(x) = 1_{\mathcal{X}}(x)$$

This is possible if and only if A_1 is the entire space X, and thus $r = 1$. Hence S is exact.

$\blacksquare$

6 THE SECOND LAW OF THERMODYNAMICS

Statistical mechanics studies macroscopic properties such as pressure and temprerature, based on microscopic concepts such as the collision between particles and the distribution of particle velocities. During the advances of this theory, new ingredients were incorporated such as entropy (thermodynamics), in this context entropy measures the complexity of the system, through the Bolzamann formula, where entropy is a function of the number of configurations that the system can have, being the most likely configuration that maximizes entropy, this is known by the name the principle of maximum entropy, even if it seems, in some obvious contexts such as in the case of water, than at room temperature The configuration that maximizes entropy is the liquid state and not the solid state, this principle is nothing more than a postulate. Subsequently, statistical models are created to be able to predict macroscopic compositions from microscopic data, three of these models are called microcanonical, canonical and grand canonical, from these models relationships between thermodynamic variables such as the ideal gas equation can be demonstrated $PV = RTN$.

During the development of chapter five we saw the entropy of Bolzmann-Gibbs and in addition to the existence of a density function that maximizes it, this density corresponds to a state of maximum entropy for thermodynamic entropy, the Proposition that guarantees the existence of this Density is called the generalized form of the microcanonic set, the particular case studied in mechanics corresponds to the Proposition 59.

During the development of thermodynamics, in the context of the industrial revolution, they postulated three laws on the behavior of thermodynamical systems, subsequently a fourth law is added, which establishes a transitivity between the equilibrium states of three bodies, this law says *If two bodies are in thermodynamic equilibrium, and the second is in thermodynamic equilibrium with a third, then the first will be in thermodynamic equilibrium with the latter* how this principle could not be deduced from the three laws and because it is such an elementary principle according to who defined it, it was decided to call it *the zero law of thermodynamics.*

We focus on the second law of thermodynamics, which basically states that,*an isolated system, that is, it does not share energy or matter with its neighborhood, where an irreversible process occurs, the entropy of the system must always increase and only remain costly if the process was reversible.*

We will use the version of the second law of thermodynamics in the book *Time's Arrow: The Origins of Thermodynamic Behavior* of Michael C. Mackey [11], in which it distinguishes from two versions of this law, a weak and a strong one, which are:
Let $S_{TD}(t)$ denote the thermodynamic entropy at time t.
Weak Form of the Second Law
$-\infty < S_{TD}(t_0) \leq S_{TD}(t) \leq 0$ for all times $t_0 < t$ and there exist a set of equilibrium entropies $\{S^*_{TD}(f)\}$ dependendet on the initial preparation f of the system such that

$$\lim_{t\to\infty} S_{TD}(t) = S^*_{TD}(f) \leq \max_f S^*_{TD}(f).$$

Thus the case system difference $\Delta S(t) = S_{TD}(t) - \max_f S^*_{TD}(f)$ satisfies $\Delta S(t) \leq 0$ and

$$\lim_{t\to\infty} \Delta S(t) \leq 0$$

In this case system entropy converges to a steady state value S^*_{TD} which may not be unique. If it is not unique it characterizes a metastable state.

Strong Form of the Second Law
$-\infty < S_{TD}(t_0) \leq S_{TD}(t) \leq 0$ for all times $t_0 < t$ and there is a unique limit S^*_{TD} (independent of the initial system preparation f) such

$$\lim_{t\to\infty} S_{TD} = S^*_{TD}$$

for all initial system preparation f. Under these circumstance,

$$\lim_{t\to\infty} \Delta S(t) = 0$$

Due to the temporal aximetry presented by this law, it was considered that this law defined a sense in the evolution of time itself. It is natural to think that such an important postulate should be demonstrable, however, the tools of statistical mechanics were not able to prove it. But it did not take long to find a theory that could with this problem, and it was when the ergodic theory, entered the scene, to study this problem we first have to translate the concepts of statistical mechanics into the concept of ergodic theory. A thermodynamical system is equivalent to a measurement space, the equilibrium states correspond to stationary densities, and instead of working with thermodynamic entropy, we consider the entropy of Bolzmann Gibbs since stationary densities would correspond to states that maximize thermodynamic entropy. The first problem we encounter is that the conception of ergodicity only guaranteed the existence of a stationary density. This is due to Theorem 39 that establishes an equivalence between ergodicity and the Cesàro convergence of the $\{P^n f\}$ where P is the operator of frobrenius perron associated with the system and $f \in \mathcal{D}$. But this does not guarantee that the system reaches the state of maximum thermodynamic entropy.

Consider an even stronger condition, consider mixing systems,remember that a system is a mixting if

$$\lim_{n\to\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B)$$

for all $A, B \in \mathcal{A}$.

now we can guarantee only the maximum entropy but locally, which corresponds to the weak version of thermodynamics.But it is enough to show the strong version. We need an even stronger condition, well consider exact system, i.e

$$\lim_{n\to\infty} \mu(S^n(A)) = 1$$

for every $A \in \mathcal{A}$ and $\mu(A) > 0$.
With this now if we can guarantee that the system reaches the state of maximum entropy, the problem was solved, or not ?, the truth is that we have complicated things more, by the following reasons, the first of them, is that at no time have we considered the dynamic variables such as temperature and pressure, and on the other hand, a necessary condition for a system to be exact, is that the function that

generates the dynamics has to be not necessarily invertible, however, the laws of thermodynamics are reversible and in general those of all physics, except the second law of thermodynamics, but we can not consider it as it is what we try to demonstrate. This generates more problems than solution. We have two options, the first and most pessimistic, is to consider that all the laws of physics are bad formulas, and that it should be reformulated into irreversible laws. The other option, more with more hope, is that the phase space, that is, the set of where we define the dynamics, is not continuous, but that this granulate, in this way, generate restrictions, which we had not considered. Another operation is that there are hidden variables that are in the worst case unacceptable to us, and that are playing a decisive role without us noticing. Whatever the path that responds to this new problem, every day around us we see the consequences of the second law of thermodynamics, because at the end of the cases, whenever we put an ice cube on the fire, we will always expect that This melts.

REFERENCES

GLOSSÁRIO

Palavra	Significado da palavra
Palavra 2	Significado da palavra 2

REFERENCES

[1] *Fundamentos de Análise Funcional* Geraldo Botelho, Daniel Pellegrino, Eduardo Teixera,, *Editorial SBM*(2012).

[2] Mixing for Markov operators, Z. Wahrscheinlichkeitstheorie,*Lin, M. 1971*, *Verw. Gebiete, 19:231-242.*(1971).

[3] *Ergodic problems of classical mechanics* V.I Arnold, A. Avez, , Princenton University (1968).

[4] *Statistical properties of Deterministic systems*, Jiu Ding, Aihui Zhou,*Springer*(2009)

[5] *"Proof of the impossibility of ergodic systems: the 1913 papers of Rosenthal and Plancherel"* S.G. Brush,, en The Kinetic Theory of Gases: An Anthology of Classical Papers with Historical Commentary (Imperial College Press, 2003) p. 505.

[6] *Foundations of ergodic theory,* Marcelo Viana, Krerley Oliveira, *Cambrindge Studies in advanced mathematics, Cambridge University Press* (2016).

[7] *Laws of Chaos: invariant measure and Dynamical System in one Dimension,* Abraham Boyarsky, Pawel Góra, *Springer Science+Business Media. LLC* New York (1997).

[8] *Linear operators. Part I: General theory,*Nelson Dunford and Jacob T. Schwartz, *Interscience Publishers*, (1958).

[9] *Non-spectral Asymptotic Analysis of One-Parameter Operator Semigroups,*Eduard Yu. Emel'yanov, Birkhäuser (2007)

[10] *Chaos, Fractals, and Noise Stochastic Aspects of Dynamics*, Andrzej Lasota, Michael C. Mackey **Springer**(1994)

[11] *Time's Arrow:The Origins of Thermodynamic Behavior*, Michael C. Mackey, *Springer Science+Business Media LLC, New York* (1992)