universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The dynamics of a taffy puller machine"

verfasst von / submitted by

Hector Homero Canales Farias

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Mathematik
Betreut von / Supervisor:	Univ.-Prof. Dr. Henk Bruin

ACKNOWLEDGEMENT

I would first like to thank my thesis advisor Prof. Dr. Henk Bruin of the Faculty of Mathematics at University of Vienna, whose door office was always open whenever I ran into a trouble spot or had a question about my work. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

Secondly, I would like to thank my previous advisor Prof. Dr. Sergio Macías Alvarez, who not only supported me during my bachelor studies but also put me in touch with my current advisor.

I would also like to acknowledge and thank my uncle Fernando Canales Lozano and my aunt Maria de la Paz Uchua de Canales who not only supported me financially during my master studies but they have always sought the welfare of our family by sharing their wisdom and advices.

Finally, I must express my very profound gratitude to my parents and brother for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Abstract

We study models of a Taffy-pulling machine described in the work of J. T. Halbert and J. A. Yorke. The action of this machine and first model is given by a 1-dimensional map constructed with a family of chaotic maps on the interval $[0, 1]$. We describe the invariant density of this model and prove some statistical properties like mixing and ergodicity. Further abstraction leads to a 2-dimensional model, an area-preserving map on a region of the plane, we represent this map as the natural extension of the 1-dimensional map. We continue the study of the 2-dimensional model and follow the work of J-L. Thiffeault to describe the map as a linear function on the torus and as a pseudo-Anosov map on the a punctured sphere.

Kurzfassung

Diese Studie befasst sich mit dem Model der Taffy-pulling Maschine von J. T. Halbert und J. A. Yorke. Jenes Model konstruiert die 1-dimensionale Funktion aus Funtkionen der Chaostheorie mit dem Intervall $[0, 1]$. Es wird die invariante Dichte dieses Models behandelt und der Beweis für einige statistischen Eigenschaften, wie zum Beispiel die Mischung und die Ergodizität, hergestellt. Die Weitere Untersuchung beschäftigt sich mit dem 2-dimensionalen Model. Dieses Model folgt der Arbeit von J-L. Thiffeault und beschreibt die Funktion als eine Pseudo-Anosov Funktion.

Contents

Chapter 1

Introduction

The aim of this work is to describe the mathematics behind the 4-rod taffy puller machine. The hearth of this work is the paper of J. T. Halbert and J. A. Yorke [3], but we will also follow very closely the combined work of F. D. Matthew and J-L. Thiffeault [2] and the paper of Thiffeault [1]. Before we start developing the mathematics of this work, we will give a short explanation of what taffy is and a brief summary of historical facts about the Taffy puller machine.

Salt Water Taffy in USA or chews in UK is a candy made of boiling sugar, butter or vegetable oil, flavourings, and colourings. The taffy is made by pulling and stretching the aforementioned candy mass. The pulling and stretching process is an important part of making taffy. Without this process, the candy would be hard and would not be as delicious as it could be. Pulling and stretching the mass allow tiny air bubbles to be captured within the candy. Those tiny air bubbles help to make the taffy soft and chewy. There are special machines that make this pulling and stretching process, which are mesmerizing to watch.

Before these special machines appeared (the first registered patent belongs to P. J. G. Firchau (19/12/1893)), taffy was made by hand. This was a difficult job to perform, since candymakers had to pull a huge mass of candy with their bare hands. They repeated the process of pulling and stretching over and over again until the taffy got the desired consistency.

The taffy candy in the USA is known by the name "Salt Water Taffy". Although the origin of the Salt water Taffy is unclear, rumor has it that everything began in New Jersey at the end of the 19th century. According to [14] and [15], the most popular story is that of a young girl who was on holidays with her family in Atlantic city. The young girl entered to a candy store, which was flooded during a storm in 1883. The girl asked to the owner if he had any taffy for sale. Since the entire merchandise of taffy was soaked with salty water, the owner in form of a joke offered the young girl some "salt water taffy". The owner's mother overheard what happened and loved the name "salt water taffy", and the name remains until today.

Chapter 2

Basic concepts

In this chapter we will introduce some of the basic concepts, definitions and results that we will use through this work.

2.0.1 Notation. *Through this work I will denote the unit interval and all our functions will be continuous unless stated otherwise.*

2.0.2 Definition. *Let $f : X \to X$ be a continuous function where X is a topological space, I and J subsets of X. We will say that I* ***f-covers*** *J if $J \subset f(I)$ and will be denoted as $I \to J$.*

2.0.3 Definition. *Let I^2 be the unit square, the space obtained by identifying the opposite edges of I^2 is the so called* ***torus****, and it will be denoted as $\mathbb{T}^2$.*

2.0.4 Observation. *$\mathbb{T}^2$ can be considered as $\mathbf{S}^1 \times \mathbf{S}^1$ or as $\mathbb{R}^2/\mathbb{Z}^2$.*

2.0.5 Definition. *Let X and Y be topological spaces, we denote the space of all continuous functions from X to Y by $\mathrm{C}(X, Y)$.*

2.0.6 Definition. *Let X and Y be topological spaces and $f \in \mathrm{C}(X, Y)$. We will call f and* ***embedding*** *if it is injective and $f^{-1} : f(X) \to X$ is continuous.*

2.0.7 Definition. *An* ***Homotopy*** *is a function $H : I \times X \to Y$ which is continuous.*

2.0.8 Definition. *Let f and $g \in \mathrm{C}(X, Y)$. We say that f is* ***homotopic*** *to g, and denoted by $f \sim g$, if and only if there exists an homotopy $H : I \times X \to Y$ with $H(0, x) = f(x)$ and $H(1, x) = g(x)$ for every $x \in X$.*

2.0.9 Observation. *In the case where f and g are embeddings, we will say that f is **isotopic** to g if the homotopy H is an **isotopy**, i.e., for each fixed $t \in I$, $H(t, x)$ gives an embedding.*

The following definitions and results concerning Thurston theory on surfaces can be found in the work of Fathi *et al.* [4]

2.0.10 Definition. *For a closed, orientable surface M ($2-$dimensional manifold) of genus $g > 1$, we will denote by S the set of isotopy classes of simple closed curves in M that are not homotopic to a point or homotopic to a boundary component of M.*

2.0.11 Definition. *Let M be a surface with $\chi(M) < 0$ (here χ is Euler characteristic). Let $\mathcal{H}$ denotes the set of all metrics on M with constant curvature $K = -1$ such that every component of the boundary of M is a geodesic. Let $\mathrm{Diff}_0(M)$ be the group of diffeomorphisms isotopic to the identity, with the C^∞ topology. The orbit space under the action of $\mathrm{Diff}_0(M)$ on $\mathcal{H}$, equipped with the quotient topology, is called **Teichmüller space** and it will be denoted by $\mathcal{T}(M) = \mathcal{T}$.*

2.0.12 Definition. *Let M be an $m-$manifold. An $p-$**dimensional foliation** $\mathcal{F}$ of M is a collection $\{M_i\}_{i \in \mathcal{I}}$, where $\mathcal{I}$ is an index set, of pairwise-disjoint, connected, immersed $p-$dimensional submanifolds of M, called the **leaves** of the foliation, such that for every $x \in M$, there is a chart (U, ϕ) with U homeomorphic to $\mathbb{R}^n$ containing x such that every leaf, M_i, meets U in either the empty set or a countable collection of subspaces whose images under ϕ in $\phi(M_i \cap U)$ are $p-$dimensional affine subspaces whose first $m-p$ coordinates are constant.*

2.0.13 Definition. *A **measured foliation** $\mathcal{F}$ on an $m-$manifold M is a foliation with isolated singularities, equipped with a positive measure on each transverse arc that is equivalent to the Lebesgue measure of a closed interval of $\mathbb{R}$, where an arc in M is a homeomorphic image of the unit interval.*

2.0.14 Definition. *Let X be a topological space, $f \in \mathrm{C}(X, X)$ and $A \subset X$. We will sat that A is **f-invariant** or just **invariant** if $f(A) \subset A$, and **completely invariant** if $f^{-1}(A) = A$.*

2.0.15 Definition. *Let M be an $m-$dimensional manifold. A compact, completely invariant subset $\Lambda \subset M$ is called **hyperbolic** if there are $\lambda \in (0, 1)$, $C > 0$, and families of subspaces $E^s(x)$, $E^u(x) \in T_xM$ with $x \in \Lambda$, such that for every $x \in \Lambda$,* *(1)* $T_xM = E^s(x) \bigoplus E^u(x)$,

(2) $\|df_x^n(v^s)\| \le C\lambda^n \|v^s\|$ *for every* $v^s \in E^s(x)$ *and* $n \ge 0$,

(3) $\|df_x^{-n}(v^u)\| \le C\lambda^n \|v^u\|$ *for every* $v^u \in E^u(x)$ *and* $n \ge 0$,

(4) $df_x E^s(x) = E^s(f(x))$ *and* $df_x E^u(x) = E^u(f(x))$.

The subspace $E^s(x)$ *(respectively,* $E^u(x)$*) is called the* ***stable (unstable) manifold*** *at* x*, and the family* $\{E^s(x)\}_{x\in\Lambda}$ *(*$\{E^u(x)\}_{x\in\Lambda}$*) is called the* ***stable (unstable) distribution*** *of* $f|_\Lambda$*.*

If $M = \Lambda$*, then* f *is called* ***Anosov diffeomorphism****.*

2.0.16 Observation. *Let* $T : \mathbb{T}^2 \to \mathbb{T}^2$ *be a linear map such that* A *has integer entries and* $\det(A) \in \{1, -1\}$*, where* A *is the matrix representation of* T*. Then* T *is and Anosov diffeomorphism.*

2.0.17 Example. *The cat map given by*

$$f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \cdot x \bmod 1,$$

the matrix has eigenvalues $\lambda_{1,2} = \frac{3\pm\sqrt{5}}{2}$*, which means that* f *is an Anosov map.*

2.0.18 Definition. *A homeomorphism* $f : M \to M$ *where* M *is a* $m-$ *manifold is called* ***pseudo-Anosov*** *if there exists a transverse pair of measured foliations on* M*,* $\mathcal{F}^s$ *(stable) and* $\mathcal{F}^u$ *(unstable), and a real number* $\lambda > 1$ *such that the foliations are preserved by* f *and their transverse measures are multiplied by* $\frac{1}{\lambda}$ *and* λ*. The number* λ *is called the* ***stretching factor*** *or* ***dilatation*** *of* f*.*

Let $\phi \in \mathrm{Diff}_0(\mathbb{T}^2)$. Up to isotopy, $\phi \in \mathrm{SL}(2,\mathbb{Z})$. There are three distinct possibilities for the eigenvalues λ_1 and λ_2 of ϕ as follows:

(a) λ_1 and λ_2 are complex. In this case, ϕ is of finite order.

(b) $\lambda_1 = \lambda_2 = 1$ (respectively $\lambda_1 = \lambda_2 = -1$). Up to a change of coordinates,

$$\phi = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \text{ respectively, } \phi = \begin{pmatrix} -1 & a \\ 0 & -1 \end{pmatrix},$$

ϕ leaves invariant a simple curve.

(c) λ_1 and λ_2 are distinct irrationals. Then ϕ is an Anosov diffeomorphism.

This analysis is generalized by Thurston to any compact surface:

2.0.19 Theorem. *Any diffeomorphism $\phi : M \to M$ is isotopic to a map $d\phi : T_xM \to T_{\phi(x)}M$ satisfying one of the following conditions:*

(1) $d\phi$ fixes an element of $\mathcal{T}(M)$ and is of finite order.

(2) $d\phi$ is "reducible", in the sense that it preserves a simple curve.

(3) there exists $\lambda > 1$ and two transverse measured foliations $\mathcal{F}^s(x)$ and $\mathcal{F}^u(x)$ such that

$$d\phi(\mathcal{F}^s(x)) = \frac{1}{\lambda_x}\mathcal{F}^s(x) \text{ and}$$

$$d\phi(\mathcal{F}^u(x)) = \lambda_x \mathcal{F}^u(x).$$

The equalities in (3) mean that the underlying foliations are equal, and the measures are scaled. Note that a pseudo- Anosov diffeomorphism satisfies condition (3).

2.1 The braid group and the Burau representation

Let X be a topological space and n a positive integer. Let $P_n(X) = X^n - \Delta$, where Δ is the big diagonal of X^n, that is the set of $n-$tuples $(x_1, \ldots, x_n)$ such that for some $i \neq j$, $x_i = x_j$. The symmetric group $Sym(n)$ acts freely on $P_n(X)$ and we define $B_n(X) = P_n(X)/Sym(n)$. One thus has a regular covering

$$\begin{tikzcd} Sym(n) \arrow[r] & P_n(X) \arrow[d] \\ & B_n(X) \end{tikzcd}$$

2.1.1 Definition. *The group $\pi_1(P_n(X))$ is called the **pure braid group** of X on n strands, and $\pi_1(B_n(X))$ is called the **braid group** of X on n strands. The classical braid group of **Artin** is the braid group $\pi_1(B_n(\mathbb{R}^2))$ and we will denote it as B_n.*

An element of B_n may be represented in the following fashion: fix a set of n distinct points $x_1, \dots, x_n$ in the interior of D^2, where D^2 is the unit disk. An element of B_n is a system of arcs in $D^2 \times I$, going from $(x_1, \dots, x_n) \times \{0\}$ to $(x_1, \dots, x_n) \times \{1\}$, transverse to every horizontal slice $D^2 \times \{t\}$. The arcs do not meet $\partial D^2 \times I$, and everything is defined up to isotopies that leave invariant the boundary of the cylinder and respect the projection $D^2 \times I \to I$. The pure braids are those for which the arc leaving $x_i \times 0$ ends at $x_i \times 1$. Figure represents an element of B_n.

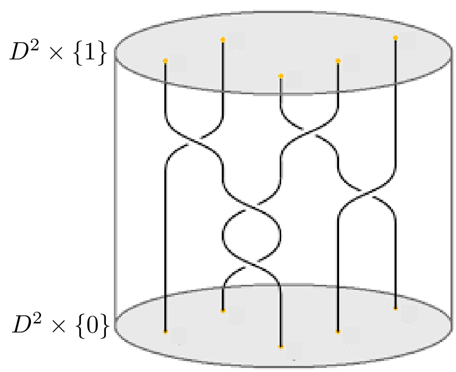


Figure 2.1: A element of B_n.

The next paragraphs are just a short summary of what can be found in [5] and [7]. Now we will define B_n in a purely algebraic way as the subgroup of *right automorphisms* α of the free group $F_n = \langle x_1, \dots, x_n \rangle$ which verify:

$$(x_i)\alpha = A_i x_{\mu_i} a_i^{-1},\ 1 \leq i \leq n,$$

$$(x_1 x_2 \cdots x_n)\alpha = x_1 x_2 \cdots x_n$$

where $(\mu_1, \dots, \mu_n)$ is a permutation of $(1, \dots, n)$ and A_i is an element of F_n. This group is of finite type with generators $\sigma_1, \dots, \sigma_{n-1}$, where

$$(x_i)\sigma_i = x_i x_{i+1} x_i^{-1},$$

$$(x_{i+1})\sigma_i = x_i,$$

$$(x_j)\sigma_i = x_j j \neq i, i+1.$$

and relations

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1},\ 1 \leq i \leq n-2,$$

$$\sigma_i\sigma_j = \sigma_j\sigma_i,\ |i-j| \geq 2.$$

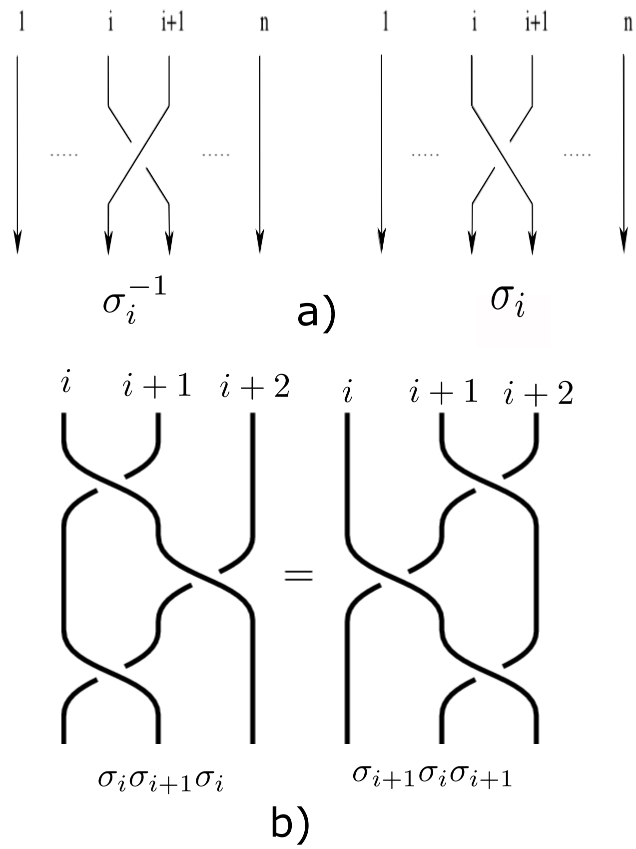


Figure 2.2: a) the generators σ_i and σ_i^{-1}, b) the relation $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$.

2.1.1 Free differential calculus

Let $\mathbb{Z}F_n$ be the ring of F_n with integer coefficients. For $j = 1, 2, \ldots, n$, there exists a well-defined map: $\dfrac{\partial}{\partial x_j} : \mathbb{Z}F_n \to \mathbb{Z}F_n$ called **free derivative** and verifying:

$$\frac{\partial}{\partial x_j}(c_1+c_2)=\frac{\partial}{\partial x_j}(c_1)+\frac{\partial}{\partial x_j}(c_2), \qquad c_1,c_2\in\mathbb{Z}F_n$$
$$\frac{\partial}{\partial x_j}(g_1g_2)=\frac{\partial}{\partial x_j}(g_1)+g_1\frac{\partial}{\partial x_j}(g_2), \qquad g_1,g_2\in F_n$$
$$\frac{\partial}{\partial x_j}(x_i)=\delta^i_j \qquad 1\le i\le n.$$

Finally, let $\langle t\rangle$ be the free group with one generator t and

$$\varphi\begin{cases}F_n\to\langle t\rangle\\ x_i\to t,\quad 1\le i\le n\end{cases}$$

φ extends into a morphism $\mathbb{Z}F_n\to\mathbb{Z}\langle t\rangle=\mathbb{Z}[t,t^{-1}]$ that we will continue to call φ.

2.1.2 Definition. *Consider an element $\alpha\in B_n$ and define the $n-$square matrix $B_\alpha=(b_{ij})$ by:*

$$b_{ij}=\varphi\left(\frac{\partial}{\partial x_j}((x_i))\alpha\right)\in\mathbb{Z}[t,t^{-1}]\quad(1\le i,j\le n).$$

*One can show the relation $B_{\alpha\beta}=B_\alpha B_\beta$. The map $B:B_n\to GL(n,\mathbb{Z}[t,t^{-1}])$ given by $\alpha\to B_\alpha$ is the **Burau representation** of the braid group B_n.*

Chapter 3

Ergodic theory concepts and results

3.0.1 Definition. *A measure μ is called* ***absolutely continuous*** *with respect to a measure λ, and denoted by $\mu \ll \lambda$, if $\lambda(A) = 0$ implies $\mu(A) = 0$. If both $\mu \ll \lambda$ and $\lambda \ll \mu$ then we will say that μ and λ are* ***equivalent****.*

3.0.2 Definition. *A measurable function $T : (X, \mathcal{A}, \lambda) \to (X', \mathcal{A}', \lambda')$ is called* ***measurable preserving*** *(m.p) if $\lambda \circ T^{-1} = \lambda'$.*

3.0.3 Definition. *$(X, \mathcal{A}, \mu, T)$ is a* ***measure preserving system*** *if T is a measure preserving function from $(X, \mathcal{A}, \mu)$ to itself. Sometimes we will say that the measure μ is* ***T-invariant***

The following Theorem is known as the Radon-Nikodym theorem and it can be found in any book of measure theory see for example [8].

3.0.4 Remark. *recall $L^1(\nu)$ is the space of all integrable functions with respect the measure ν.*

3.0.5 Theorem. *If μ is a probability measure and $\mu \ll \nu$ then there is a function $h \in L^1(\nu)$ such that $\mu(A) = \int_A h(x) d\nu(x)$ for every measurable set A, h is called* ***Radon-Nikodym density****. Sometimes we use the notation $h = \frac{d\mu}{d\nu}$.*

3.0.6 Proposition. *Let $T : U \to U$ be (piecewise) differentiable function where $U \subset \mathbb{R}^n$, and $\mu \ll \lambda$ where λ is the Lebesgue measure. Then μ is*

T-invariant if and only if its density $h = \frac{d\mu}{dx}$ satisfies

$$\hat{T}h(x) = \sum_{y \in T^{-1}(x)} \frac{h(y)}{|det(DT(y))|} = h(x) \tag{3.1}$$

for every x. We will called $\hat{T}$ the ***transfer operator****.*

Proof. the T-invariance means that $d\mu(x) = d\mu(T^{-1}(x))$, but we need to beware that T^{-1} is multivalued. So we will split U into pieces U_n such that the restrictions $T_n := T|_{U_n}$ are diffeomorphism (onto their images) and write $y_n = T_n^{-1}(x) = T^{-1}(x) \cap U_n$. Then using change of coordinates we obtain

$$h(x)dx = d\mu(x) = d\mu(T^{-1}(x)) = \sum_n d\mu \circ T_n^{-1}(x)$$

$$= \sum_n h(y_n)|det(DT_n^{-1})(x)|dy_n = \sum_n \frac{h(y_n)}{|det(DT(y_n))|} dy_n,$$

and the statement follows. Conversely, if (2.1) holds, then the above computation gives $d\mu(x) = d\mu \circ T^{-1}(x)$, which is the required invariance. □

3.0.7 Definition. *Two measure preserving system $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$ are called* ***isomorphic*** *if there are $X' \in \mathcal{A}$, $Y' \in \mathcal{B}$ and $\phi : X' \to Y'$ such that*

1.- $\mu(X') = 1$, $\nu(Y') = 1$;

2.- $\phi : X' \to Y'$ is a bi-measurable bijection;

3.- ϕ is measure preserving; i.e., $\mu(\phi^{-1}(B)) = \nu(B)$ for every $B \in \mathcal{B}$;

4.- $\phi \circ T = S \circ \phi$.

3.0.8 Definition. *Let $(Y, \mathcal{B}, \nu, S)$ be a measure preserving system. A system $(X, \mathcal{A}, \mu, T)$ is called a* ***natural extension*** *of $(Y, \mathcal{B}, \nu, S)$ if there are $X' \in \mathcal{A}$, $Y' \in \mathcal{B}$ and $\phi : X' \to Y'$ such that*

1.- $\mu(X') = 1$, $\nu(Y') = 1$;

2.- $\phi : X' \to Y'$ is a measurable surjection;

3.- ϕ is measure preserving; i.e., $\mu(\phi^{-1}(B)) = \nu(B)$ for every $B \in \mathcal{B}$;

4.- $\phi \circ T = S \circ \phi$;

5.- $T : X' \to X'$ is invertible

We will use the following kinds of convergence in $\mathrm{L}_p(\lambda)$ spaces

3.0.9 Definition. *1) Norm or strong convergence: $f_n \to f$ in $\mathrm{L}_p(\lambda)$ if and only if*

$$\|f_n - f\|_p \to 0 \text{ as } n \to \infty, \text{ where } \|f\|_p = \left(\int |f(x)|^p d\lambda \right)^{\frac{1}{p}};$$

2) Weak convergence: $f_n \to f$ weakly in $\mathrm{L}_p(\lambda)$ $1 \le p < \infty$ if and only if for every $g \in \mathrm{L}_q(\lambda)$, $\int f_n g d\lambda \to \int f g d\lambda$, where $\frac{1}{p} + \frac{1}{q} = 1$;

3 almost everywhere convergence (a.e.): $f_n \to f$ if and only if $f_n(x) \to f(x)$ for almost all $x \in X$.

The following result is known as Hölder's inequality and we will make use of it later. For a proof please consult any book in measure theory like [8].

3.0.10 Theorem. *Let $(X, \mathcal{A}, \lambda)$ be a measure space and $p,\ q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for every $f \in \mathrm{L}_p(\lambda)$ and $q \in \mathrm{L}_q(\lambda)$ we have*

$$\|f \cdot g\|_1 \le \|f\|_p \|g\|_q .$$

Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system. If $T^{-1}A = A$ for some $A \in \mathcal{A}$, then $T^{-1}A^{\mathsf{c}} = A^{\mathsf{c}}$ and the study of T split into two parts: $T|_A$ and $T|_{A^{\mathsf{c}}}$. It is useful to have a concept of *indecomposability* for measure preserving functions, so that if T has this indecomposable property then the study of T cannot be split into separate parts. This property is called *ergodicity*.

3.0.11 Definition. *We call a measure preserving transformation $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ **ergodic** if for any $A \in \mathcal{A}$, such that $T^{-1}A = A$, $\mu(A) = 0$ or $\mu(A^{\mathsf{c}}) = 0$. We define $\mathcal{I}(T) = \{A \in \mathcal{A};\ T^{-1}A = A\}$.*

The following results are well known and can be found in any book of ergodic theory. We will not give a proof, for more details please see for example [9] or [11]

3.0.12 Proposition. *Let $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ be a measure preserving function. Then the following are equivalent:*

1) T is ergodic

2) If $f \in \mathrm{L}_1(\mu)$ and $f \circ T = f$ a.e., then f is constant a.e.

Let T be like above and $A \in \mathcal{A}$. For $x \in X$, a question of physical interest is: With what frequency do the points of the orbit $\{x, Tx, \ldots\}$ occur in the set A?

Clearly, $T^i x \in A$ if and only if $\mathbb{1}_A(T^i x) = 1$. Thus, the number of points of $\{x, Tx, \ldots, T^{n-1}x\}$ in A is equal to $\sum_{k=0}^{n-1} \mathbb{1}_A(T^k x)$, and the relative frequency is $\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A(T^k x)$. So we define the **Birkhoff sum** of $f \in \mathrm{L}_1(\mu)$ as $\mathrm{S}_n(f) = \sum_{k=0}^{n-1} f \circ T^k$. The following Theorem is known as the Birkhoff ergodic Theorem.

3.0.13 Theorem. *Let $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ be measure preserving system. Then for every $f \in \mathrm{L}_1(\mu)$,*

$$\frac{1}{n}\,\mathrm{S}_n(f) \to f^* \ \textit{a.e.},\ f^* \in \mathrm{L}_1(\mu).$$

Furthermore, $f^* \circ T = f^*$ a.e. and if $\mu(X) < \infty$ and T is ergodic, then $f^* = \int_X f d\mu$.

The next Theorem is known as the $\mathrm{L}_p(\mu)$ ergodic Theorem.

3.0.14 Theorem. *Let $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ be measure preserving system. If T is ergodic, and $f \in \mathrm{L}_p(\mu)$ for some $1 \leq p < \infty$ then there exists $f^* \in \mathrm{L}_p(\mu)$ with $f^* \circ T = f^*$ a.e. such that*

$$\left\| \frac{1}{n}\,\mathrm{S}_n(f) - f^* \right\|_p \to 0 \ \textit{as}\ n \to \infty.$$

3.0.15 Proposition. *Let $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ be measure preserving system on a normalized measure space. Then T is ergodic if and only if for all $A,\ B \in \mathcal{A}$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) \to \mu(A)\mu(B).$$

3.0.16 Definition. *We say that measure preserving function $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ is* ***weakly mixing*** *if for all A, $B \in \mathcal{A}$*

$$\frac{1}{n}\sum_{k=0}^{n-1} |\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)| \to 0 \text{ as } n \to \infty.$$

T is ***strongly mixing or mixing*** *if for all A, $B \in \mathcal{A}$,*

$$\mu(T^{-n}(A) \cap B) \to \mu(A)\mu(B) \text{ as } n \to \infty$$

3.0.17 Proposition. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving dynamical system. Then, T is mixing if and only if*

$$\int_X f \cdot g \circ T^n d\mu \to \int_X f d\mu \int_X g d\mu,$$

for every $f \in \mathrm{L}_1(\mu)$ and $g \in \mathrm{L}_\infty(\mu)$.

3.0.18 Definition. *We say that measure preserving function $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ is* ***exact*** *if the tail σ-algebra $\mathcal{T}(\mathcal{A}) = \bigcap_{k\geq 0} T^{-k}\mathcal{A}$ consists of the sets of measure 0 or 1;i.e., for every $A \in \mathcal{T}(\mathcal{A})$ we have that $\mu(A) = 0$ or $\mu(A^{\mathsf{c}}) = 0$.*

3.0.19 Proposition. *We have the following implications:*
mixing $\implies$ weakly mixing $\implies$ ergodic.

3.1 Functions of bounded variation

Let $[a, b] \subset \mathbb{R}$ be a bounded interval and let λ denote Lebesgue measure on $[a, b]$. For any sequence of points $a = a_0 < a_1 < \ldots < a_n = b$, $n \geq 1$, we define a partition $\mathcal{P} = \{I_i = [a_{i-1}, a_i);\ i = 1, \ldots, n\}$ of $[a, b]$. The points $\{a_0, \ldots, a_n\}$ are called *endpoints of the partition $\mathcal{P}$*.

3.1.1 Definition. *Let $f : [a, b] \to \mathbb{R}$ and let $\mathcal{P}$ be a partition of $[a, b]$. If there exists a positive number M such that*

$$\sum_{k=1}^{n} |f(a_k) - f(a_{k-1})| \leq M$$

for all partitions $\mathcal{P}$, then T is said to be of ***bounded variation*** *on $[a, b]$.*

If f is increasing or if it satisfies the Lipschitz condition

$$|f(x) - f(y)| < K|x - y|,$$

then it is of bounded variation.

3.1.2 Definition. *Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation. The number*

$$\mathrm{V}_{[a,b]}\, f = \sup_{\mathcal{P}} \left\{ \sum_{k=1}^{n} |f(a_k) - f(a_{k-1})| \right\}$$

*is called the **total variation** of f on $[a, b]$.*

3.1.3 Proposition. *Let f be a function of bounded variation such that $\|1\|_f < \infty$. Then*

$$|f(x)| \le \mathrm{V}_{[a,b]}\, f + \frac{\|f\|_1}{b-a}$$

for all $x \in [a, b]$.

Proof. We claim that there exists $y \in [a, b]$ such that $|f(y)| \le \frac{\|f\|_1}{b-a}$. If not, then for any $x \in [a, b]$

$$(b - a)|f(x)| > \|f\|_1 .$$

Hence,

$$\|f\|_1 = \int_a^b |f(x)| d\lambda(x) > \int_a^b \frac{\|f\|_1}{b-a} d\lambda(x) = \|f\|_1$$

and we have a contradiction.

Since $|f(x)| \le |f(x) - f(y)| + |f(y)|$ we have

$$|f(x)| \le \mathrm{V}_{[a,b]}\, f + \frac{\|f\|_1}{b-a}.$$

□

We present a result due to E. Helly that we will use later. For a proof of the theorem please see [10] page 220

3.1.4 Theorem. *Let an infinite family of functions $\mathcal{F}$ be defined on an interval $[a, b]$. If all functions of the family and the variation of all functions of the family are bounded by a single number, i.e., $|f(x)| \le K$, $\mathrm{V}_{[a,b]}\, f \le K$ for every $f \in \mathcal{F}$, then there exists a sequence $\{f_n\}_{n \ge 1} \subset \mathcal{F}$ that converges pointwise to some function f^* of bounded variation and $\mathrm{V}_{[a,b]}\, f^* \le K$.*

We now make the space of functions of bounded variation into a Banach space. Let

$$\mathrm{BV}([a,b]) = \{f \in \mathrm{L}_1(\lambda);\ \inf_{f_1=f \text{ a.e}} \mathrm{V}_{[a,b]}\, f_1 < \infty\}.$$

Note that the infimum is taken over all functions a.e equal to f. For example the function

$$f(x) = \begin{cases} n, \text{ if } x = \frac{1}{n},\ n = 1, 2 \ldots \\ 0, \text{ otherwise,} \end{cases}$$

clearly has infinite variation but $f \in \mathrm{BV}([0,1])$ since $f_1 = 0$ is a.e equal to f and $\mathrm{V}_{[0,1]}\, f_1 = 0$.

We define a norm on $\mathrm{BV}([a,b])$ as follows: For $f \in \mathrm{BV}([a,b])$,

$$\|f\|_{\mathrm{BV}} = \|f\|_1 + \inf_{f_1=f \text{ a.e.}} \mathrm{V}_{[a,b]}\, f_1$$

Without $\mathrm{L}_1(\lambda)$-norm, $\|\cdot\|_{\mathrm{BV}}$ would not be a norm, since a function that is not 0 λ a.e. could have 0 variation. Now we give some properties of the space $\mathrm{BV}([a,b])$, all proofs can be found in [9].

3.1.5 Proposition. $\mathrm{BV}([a,b])$ *is dense in* $\mathrm{L}_1(\lambda)$.

Proof. Since $\mathrm{C}^1([a,b])$ is dense in $\mathrm{L}_1(\lambda)$ and since $\mathrm{C}^1([a,b]) \subset \mathrm{BV}([a,b])$ the result follows. □

The following Theorem make a connection between weakly convergence and norm convergence in $\mathrm{L}_1(\lambda)$. Recall Definition 3.0.9, for the definition of weakly and norm convergence.

3.1.6 Theorem. *If* $f_n \to f$ *weakly in* $\mathrm{L}_1(\lambda)$ *and almost everywhere, then* $f_n \to f$ *strongly in* $\mathrm{L}_1(\lambda)$

3.1.7 Proposition. *A bounded set in* $\mathrm{BV}([a,b])$ *is relatively compact in* $\mathrm{L}_1(\lambda)$*;i.e., every sequence in such bounded set has a convergent subsequence*

Proof. If $\{f_i\}_{i\in\mathfrak{I}} \subset \mathrm{BV}([a,b])$ is bounded, there is $K_1 < \infty$ such that $\|f_i\|_{\mathrm{BV}} \le K_1$ for every $i \in \mathfrak{I}$.

From the definition of $\|\cdot\|_{\mathrm{BV}}$ if follows that $\{f_i\}_{i\in\mathfrak{I}}$ is uniformly bounded; *i.e.*, there is $K_2 < \infty$ such that $|f_i(x)| \le K_2$ for every $i \in \mathfrak{I}$. Let $K =$

$\max(K_1, K_2)$. By Helly's Theorem 3.1.4, there exists a subsequence $\{f_{i_j}\}$ such that $f_{i_j} \to f^*$ everywhere. Since $\{f_i\}_{i \in \mathcal{I}}$ is weakly compact (it is uniformly bounded) and $f_{i_j} \to f^*$, Theorem 3.1.6 implies that $f_{i_j} \to f^*$ in $\mathrm{L}_1(\lambda)$. Hence $\{f_i\}_{i \in \mathcal{I}}$ is strongly compact in $\mathrm{L}_1(\lambda)$. □

3.1.8 Proposition. *If* $\mathrm{V}_{[a,b]} f_n \leq K$ *for all* n *and* $f_n \to f$ *in* $\mathrm{L}_1(\lambda)$*, then* $\mathrm{V}_{[a,b]} f \leq K$

Proof. Since $f_n \to f$ in $\mathrm{L}_1(\lambda)$, we can assume that some subsequence $f_{n_k} \to f$ everywhere after changing the functions on a set of measure 0. Consider the partition $a = a_0 < \ldots < a_n = b$. Then we have

$$\sum_{i=1}^{n} |f_{n_k}(a_i) - f_{n_k}(a_{i-1})| \leq K, \; k = 1, 2, \ldots$$

Taking limit as k goes to infinity, we obtain

$$\sum_{i=1}^{n} |f(a_i) - f(a_{i-1})| \leq K.$$

Since the partition was arbitrary, we have $\mathrm{V}_{[a,b]} f \leq K$. □

3.2 The Frobenius-Perron operator

The following definitions, results and ideas are taken from the work of Boyarsky and Góra, for more details please consult [9].

Let's start with some motivation. Let X be a random variable on the space $I = [a, b]$ having the probability density function f. Then for any measurable set $A \subset I$,

$$\mathrm{Prob}(X \in A) = \int_A f d\lambda,$$

where λ is the normalized Lebesgue measure on I. Let $T : I \to I$ be a continuous and non-singular map (for this part of the work we will always assume that our functions are non-singular unless otherwise stated). Then $T(X)$ is also a random variable and it is reasonable to ask: What is the probability density function of $T(X)$? We write

$$\mathrm{Prob}(T(X) \in A) = \mathrm{Prob}(X \in T^{-1}A) = \int_{T^{-1}A} f d\lambda.$$

To obtain a probability density function for $T(X)$, we have to write the last integral as

$$\int_A \phi d\lambda,$$

for some function ϕ. We note that if such function exists it will depend on f and T. Define

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

where $f \in \mathrm{L}_1(\lambda)$ and A is an arbitrary measurable set. If $\lambda(A) = 0$, then $\lambda(T^{-1}A) = 0$, which in turn implies that $\mu(A) = 0$. Therefore $\mu \ll \lambda$, by the Radon-Nikodym Theorem 3.0.5, there exists a $\phi \in \mathrm{L}_1(\lambda)$ such that for all measurable sets A we have

$$\mu(A) = \int_A \phi d\lambda,$$

ϕ is unique up to sets of measure zero and depends on T and f. Set $\mathrm{P}_T f = \phi$. Thus, the probability density function f has been transformed to a new probability density function $\mathrm{P}_T f$. P_T is referred to as the **Frobenius-Perron operator** associated with T, which leads us to the next definition

3.2.1 Definition. *Let $(I, \mathcal{A}, \lambda, T)$ with λ the normalized Lebesgue measure on I. We define the **Frobenius-Perron operator** $\mathrm{P}_T : \mathrm{L}_1(\lambda) \to \mathrm{L}_1(\lambda)$ as follows: For any $f \in \mathrm{L}_1(\lambda)$, $\mathrm{P}_T f$ is the unique (up to a.e equivalence) function in $\mathrm{L}_1(\lambda)$ such that for any $A \in \mathcal{A}$*

$$\int_A \mathrm{P}_T f d\lambda = \int_{T^{-1}A} f d\lambda.$$

The validity of this definition, i.e., the existence and the uniqueness of $\mathrm{P}_T f$, follows by the Radon-Nikodym Theorem 3.0.5.

The following Theorem states some properties of the Forbenius-Perron operator $\mathrm{P}_T f$. Before going to the Theorem recall that the **Koopman operator** $\mathrm{U}_T : \mathrm{L}_\infty(\lambda) \to \mathrm{L}_\infty(\lambda)$ is defined by $\mathrm{U}_T g = g \circ T$ and that for $f \in \mathrm{L}_1(\lambda)$, $g \in \mathrm{L}_\infty(\lambda)$, we denote $\int_I fg d\lambda$ by $\langle f, g \rangle$.

3.2.2 Theorem. *Let $(I, \mathcal{A}, \lambda, T)$ and $\mathrm{P}_T : \mathrm{L}_1(\lambda) \to \mathrm{L}_1(\lambda)$ be the Frobenius-Perron operator then we have that*

1) P_T *is a positive (if* $f \geq 0$*, then* $\mathrm{P}_T f \geq 0$*) and linear operator;*

2) Preservation of Integrals: $\int_I \mathrm{P}_T f d\lambda = \int_I f d\lambda$

3) Contraction property; i.e., $\|\mathrm{P}_T f\|_1 \leq \|f\|_1$ *for any* $f \in \mathrm{L}_1(\lambda)$*;*

4) Composition property: Let $S : I \to I$*. Then* $\mathrm{P}_{T\circ S} f = \mathrm{P}_T \circ \mathrm{P}_S f$*. In particular,* $\mathrm{P}_{T^n} f = \mathrm{P}_T^n f$*;*

5) Adjoint Property. If $f \in \mathrm{L}_1(\lambda)$ *and* $g \in \mathrm{L}_\infty(\lambda)$*, then* $\langle \mathrm{P}_T f, g\rangle = \langle f, \mathrm{U}_T g\rangle$*;*

6) $\mathrm{P}_T f^* = f^*$ *a.e if and only if the measure* $\mu = f^* \cdot \lambda$*, defined by* $\mu(A) = \int_A f^* d\lambda$*, is* T*-invariant for all* $A \in \mathcal{A}$*, where* $f^* \geq 0$*,* $f^* \in \mathrm{L}_1(\lambda)$ *and* $\|f^*\|_1 = 1$*.*

Proof. 1) Let $f \in \mathrm{L}_1(\lambda)$ and assume $f \geq 0$, we will show that $\mathrm{P}_T f \geq 0$. For A measurable,

$$\int_A \mathrm{P}_T f d\lambda = \int_{T^{-1}A} f d\lambda \geq 0.$$

Since A was arbitrary, it follows that $\mathrm{P}_T f \geq 0$.

Now let α and β be constants. Then if f, $g \in \mathrm{L}_1(\lambda)$,

$$\begin{aligned}\int_A \mathrm{P}_T(\alpha f + \beta g) d\lambda &= \int_{T^{-1}A} (\alpha f + \beta g) d\lambda = \int_{T^{-1}A} \alpha f d\lambda + \int_{T^{-1}A} \beta g d\lambda \\ &= \alpha \int_A \mathrm{P}_T f d\lambda + \beta \int_A \mathrm{P}_T g d\lambda \\ &= \int_A (\alpha \mathrm{P}_T f + \beta \mathrm{P}_T g) d\lambda.\end{aligned}$$

Since is true for any measurable A, the conclusion follows.

2) Since

$$\int_I \mathrm{P}_T f d\lambda = \int_{T^{-1}I} f d\lambda = \int_I f d\lambda,$$

the result follows.

3) Let $f \in \mathrm{L}_1(\lambda)$, recall that $f = f^+ - f^-$ with f^+ and $f^- \in \mathrm{L}_1(\lambda)$ and $|f| = f^+ + f^-$. By linearity of the operator we have

$$\mathrm{P}_T f = \mathrm{P}_T f^+ - \mathrm{P}_T f^-.$$

Hence,

$$|\mathrm{P}_T f| \leq |\mathrm{P}_T f^+| + |\mathrm{P}_T f^-| = \mathrm{P}_T f^+ + \mathrm{P}_T f^- = \mathrm{P}_T |f|,$$

and

$$\|\mathrm{P}_T f\|_1 \leq \int_I |\mathrm{P}_T f| d\lambda \leq \int_I \mathrm{P}_T |f| = \int_I |f| d\lambda = \|f\|_1 .$$

4) Let $f \in \mathrm{L}_1(\lambda)$ and define the measure $\mu(A) = \int_{(T \circ S)^{-1}(A)} f d\lambda$.

Since T and S are nonsingular, $\mu \ll \lambda$ and there exists a function $\mathrm{P}_{T \circ S} f$ such that $\mu(A) = \int_A \mathrm{P}_{T \circ S} f d\lambda$. Now

$$\int_A \mathrm{P}_T \mathrm{P}_S f d\lambda = \int_{S^{-1}(T^{-1}A)} f d\lambda$$

and

$$\int_A \mathrm{P}_T \mathrm{P}_S f d\lambda = \int_{T^{-1}A} \mathrm{P}_S f d\lambda = \int_{S^{-1}(T^{-1}A)} f d\lambda.$$

Hence $\mathrm{P}_{T \circ S} f = \mathrm{P}_T \mathrm{P}_S f$ a.e. By induction, it follows that $\mathrm{P}_{T^n} f = \mathrm{P}_T^n f$.

5) Let A be a measurable set and $g = \mathbb{1}_A$. We will show that

$$\int_I (\mathrm{P}_T f) \cdot g d\lambda = \int_I f \cdot \mathrm{U}_T g d\lambda.$$

The left hand side of the above equality is

$$\int_A \mathrm{P}_T f d\lambda = \int_{T^{-1}A} f d\lambda$$

and the right hand side is

$$\int_I f \cdot (\mathbb{1}_A \circ T) d\lambda = \int_I f \cdot \mathbb{1}_{T^{-1}A} d\lambda = \int_{T^{-1}A} f d\lambda.$$

Hence the result it is verified for characteristic functions. Since the linear combination of characteristic functions are dense in $L_\infty(\lambda)$, the result holds for any $f \in L_1(\lambda)$.

6) Assume that $\mu(T^{-1}A) = \mu(A)$ for any measurable set A. Then

$$\int_{T^{-1}A} f^* d\lambda = \int_A f^* d\lambda$$

and therefore

$$\int_A \mathrm{P}_T\, f^* d\lambda = \int_A f^* d\lambda.$$

Since A was arbitrary, $\mathrm{P}_T\, f^* = f^*$ a.e.

Assume $\mathrm{P}_T\, f^* = f^*$ a.e. Then

$$\int_A \mathrm{P}_T\, f^* d\lambda = \int_A f^* d\lambda = \mu(A)$$

by definition,

$$\int_A \mathrm{P}_T\, f^* d\lambda = \int_{T^{-1}A} f^* d\lambda = \mu(T^{-1}A)$$

and so $\mu(T^{-1}A) = \mu(A)$.

□

Let $\mathcal{T}(I)$ de class of functions $T : I \to I$ that satisfy the following conditions:

1 T is *piecewise expanding*, *i.e*, there exists a partition $\mathcal{P} = \{I_i = [a_{i-1}, a_i]$ $i = 1, \dots, q\}$ of I such that $T|_{I_i}$ is $\mathbf{C}^1$, and $|T_i'(x)| \geq \alpha > 1$ for any i and for all $x \in (a_{i-1}, a_i)$, where $T_i = T|_{I_i}$;

2 $g(x) = \frac{1}{|T'(x)|}$ is a function of bounded variation, where $T'(x)$ is the appropriate one-sided derivative at the endpoints of $\mathcal{P}$.

For each $n \geq 1$, we define the partition $\mathcal{P}^{(n)}$ as follows:

$$\mathcal{P}^{(n)} = \bigwedge_{k=0}^{n-1} T^{-k}(\mathcal{P}) = \{I_{i_0} \cap T^{-1} I_{i_1} \cap \dots \cap T^{-n+1} I_{i_n};\ I_{i_j} \in \mathcal{P}\}.$$

It follows that T^n is piecewise expanding on $\mathcal{P}^n$ if T is piecewise expanding on $\mathcal{P}$.

3.2.3 Example. *The doubling map* $T : [0,1] \to [0,1]$ *given by* $Tx = 2x \mod 1$. *Figure 2.3 shows the graph of the doubling map.*

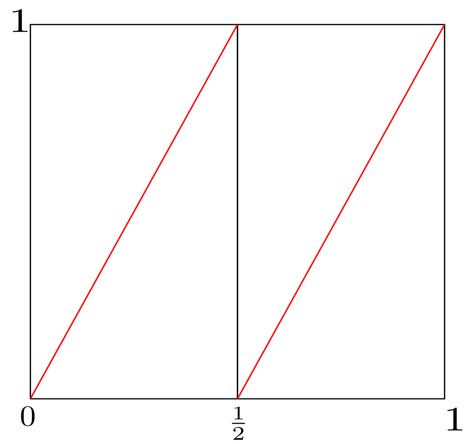


Figure 3.1: Doubling map

3.2.4 Theorem. *Let* $T \in \mathcal{T}(I)$. *Then it admits an absolutely continuous invariant measure whose density is of bounded variation.*

Consider two complex Banch spaces $(\mathcal{X}, |||\cdot|||)$ and $(\mathcal{L}, \|\cdot\|)$ with $\mathcal{X} \subset \mathcal{L}$. A linear operator $\mathrm{P} : \mathcal{X} \to \mathcal{X}$ is bounded with respect to both $|||\cdot|||$ and $\|\cdot\|_{\mathcal{X}}$, where

$$\|\mathrm{P}\|_{\mathcal{X}} = \sup\left\{\frac{\|\mathrm{P}\, f\|}{\|f\|_{\mathcal{X}}}, f \in \mathcal{X}, f \neq 0\right\}.$$

Assume that

1) If $f_n \in \mathcal{X}$, $f \in \mathcal{L}$, $\lim_{n\to\infty} \|f_n - f\| = 0$, and $|||f_n||| \leq C$ for all n, then $f \in \mathcal{X}$ and $|||f||| \leq C$;

2) $H = \sup_{n\geq 0} \|\mathrm{P}^n\|_{\mathcal{X}} < \infty$;

3) There exists $k \geq 1$, $0 < r < 1$, and $R < \infty$ such that for $f \in \mathcal{X}$,

$$\left|\left|\left|\mathrm{P}^k\, f\right|\right|\right| \leq r|||f||| + R\,\|f\|\,; \tag{3.2}$$

4) If $\mathcal{X}_1$ is a bounded subset of $\mathcal{X}$, then the closure of $\mathrm{P}^k\, \mathcal{X}_1$ is compact in $\mathcal{L}$.

For any complex number η, let

$$D(\eta) = \{f \in \mathcal{X};\ \mathrm{P}\, f = \eta f, f \neq 0\},$$

that is: η is an eigenvalue of P if and only if $D(\eta) \neq \emptyset$.

We state the Ionescu-Tulcea and Marinescu Theorem

3.2.5 Theorem. *Under conditions* $1) - 4)$, *the intersection of the spectrum of* P *with the unit circle is a set* G *of eigenvalues of* P *of modulus* 1 *which has only a finite number of elements. For each* $\eta \in G$, $D(\eta)$ *is finite-dimensional. Furthermore, there exist bounded linear operators* Q_η *and* S *on* $\mathcal{X}$ *such that*

$$\mathrm{P}^n = \sum_{\eta \in G} \eta^n\, \mathrm{Q}_\eta + \mathrm{S}^n, \tag{3.3}$$

$$\mathrm{Q}_\eta\, \mathrm{Q}_{\eta'} = 0 \text{ if } \eta \neq \eta',\ \mathrm{Q}_\eta^2 = \mathrm{Q}_\eta, \tag{3.4}$$

$$\mathrm{Q}_\eta\, \mathrm{S} = \mathrm{S}\, \mathrm{Q}_\eta = 0, \tag{3.5}$$

$$\mathrm{Q}_\eta\, \mathcal{X} = \mathrm{Cl}(D(\eta)), \tag{3.6}$$

$$\rho(\mathrm{S}) < 1, \tag{3.7}$$

where $\rho(\mathrm{S}) = \lim_{n\to\infty} \|\mathrm{S}^n\|^{\frac{1}{n}}$ *is the spectral radius of* S.

The properties $(2.3) - (2.7)$ of P constitute one of equivalent definitions of a **quasi-compact operator**.

We shall use this theorem in the following setting: $\mathcal{L} = \mathrm{L}_1(\lambda)$, $\mathcal{X} = \mathrm{BV}(I)$ and $\mathrm{P} = \mathrm{P}_T$ is the Frobenius-Perron operator.

The following Lemma will help us in the proof of Theorem 3.2.7. The proof can be found in [9], page 100.

3.2.6 Lemma. *Let* $T \in \mathcal{T}(I)$. *Then there exists constants* $0 < r < 1$, $C > 0$ *and* $R > 0$ *such that for any* $f \in \mathrm{BV}(I)$ *and any* $n \geq 1$,

$$\|\mathrm{P}_T^n\, f\|_{\mathrm{BV}} \leq C r^n\, \|f\|_{\mathrm{BV}} + R\, \|f\|_1$$

3.2.7 Theorem. *Let* $T \in \mathcal{T}(I)$. *Then the Frobenius-Perron operator* P_T *is quasi-compact on the space* $\mathrm{BV}(I)$. *In particular,*

1) *T has a finite number of ergodic absolutely continuous invariant measures: $\mu_1 = f_1\lambda, \cdots, \mu_n = f_n\lambda$, where $f_1, \cdots, f_n \in \mathrm{BV}(I)$.*

2) *For each $1 \leq i \leq n$, there exists a finite collection of disjoint sets $\{a_j^{(i)}\}_{j=1}^{n(i)}$ such that $\sup \mu_i = \bigcup_{k=1}^{n(i)} A_k^{(i)}$, and such that for $f_j^{(i)} = f_i \mathbb{1}_{A_j^{(i)}}$, $j = 1, 2, \cdots, n(i)$,*

$$\mathrm{P}_T\, f_j^{(i)} = f_{j+1}^{(i)}\ (n(i) + 1 = 1),$$

and the dynamical system $(A_j^{(i)}, T^{n(i)}, f_j^{(i)}\lambda)$ is exact.

Proof. We will proof that all condition $1) - 4)$ of Theorem 3.2.5 are satisfied.

1) Let $\{f_n\} \subset \mathrm{BV}(I)$, $\|f\|_{\mathrm{BV}} \leq C$ and $\|f_n - f\|_1 \to 0$ as $n \to \infty$, where $f \in \mathrm{L}_1(\lambda)$. We must show that $f \in \mathrm{BV}(I)$ and $\|f\|_{\mathrm{BV}} \leq C$. For this, let $\epsilon > 0$. Since $f_n \to f$ in $\mathrm{L}_1(\lambda)$, for n sufficiently large,

$$\|f\|_1 - \epsilon \leq \|f_n\|_1 \leq \|f\|_1 + \epsilon.$$

Then,

$$\mathrm{V}_I\, f_n = \|f_n\|_{\mathrm{BV}} - \|f_n\|_1 \leq C - \|f\|_1 + \epsilon.$$

By Proposition 3.1.3,

$$\|f_n\|_\infty \leq \mathrm{V}_I\, f_n + \frac{1}{\lambda(I)} \int_I |f_n| d\lambda \leq C - \|f\|_1 + \frac{\|f\|_1 + \epsilon}{\lambda(I)} + \epsilon.$$

Thus, we can apply Helly's Theorem (Theorem 3.1.4) to obtain a subsequence $\{f_{n_k}\}$ and a function f^* such that $f_{n_k}(x) \to f^*(x)$ for any $x \in I$. But $f_{n_k} \to f$ in $\mathrm{L}_1(\lambda)$. So $f^* = f$ a.e. For any partition $a = a_0 < \ldots < a_m = b$, we have

$$\sum_{i=1}^{m} |f^*(a_i) - f^*(a_{i-1})| = \lim_{k\to\infty} \sum_{i=1}^{m} |f_{n_k}(a_i) - f_{n_k}(a_{i-1})| \leq C - \|f\|_1 + \epsilon.$$

Thus, $\mathrm{V}_I\, f^* \leq C - \|f\|_1 + \epsilon$ and

$$\|f^*\|_{\mathrm{BV}} \leq C - \|f\|_1 + \epsilon + \|f\|_1 = C + \epsilon.$$

Since ϵ was arbitrary 1) is proven.

2) Since $\|\mathrm{P}_T f\|_1 \leq 1$, the second condition is satisfied.

3) Note that by Lemma 3.2.6, $k_0 \geq 1$ can be chosen so that $Cr^{k_0} < 1$.

4) Let $B \subset \mathrm{BV}(I)$ be bounded. Using the same argument as in 1), we can show that $\bar{B}$ is compact in $\mathrm{L}_1(\lambda)$. Since P_T is a continuous operator, $\overline{\mathrm{P}_T B} = \mathrm{P}_T \bar{B}$ is also compact in $\mathrm{L}_1(\lambda)$.

$\square$

3.2.8 Remark. *We invoke Theorem 3.2.5, which provides a spectral representation of the operator* $\mathrm{P}_T : \mathrm{BV}(I) \to \mathrm{BV}(I)$*, for the particular case that we will work with, we have an invariant density* ρ*, then we know that* $\mathrm{P}_T \rho = \rho$*. Set* $\mathrm{Q}\, u = \int_I u \cdot \rho d\lambda$ *and* $\mathrm{S}\, u = \mathrm{P}_T u - \mathrm{Q}\, u$*, then* $\mathrm{P}_T u = \mathrm{Q}\, u + \mathrm{S}\, u$ *for every* $u \in \mathrm{BV}(I)$*. We have that* Q *is the projection onto the space* $\mathcal{H} = \{c \cdot \rho |\ c \text{ is a constant}\}$ *and both operators* Q *and* S *fulfil the properties stated in Theorem 3.2.5. From this properties we can deduce*

$$\mathrm{P}_T^n u = \mathrm{Q}\, u + \mathrm{S}^n u,$$

for every $n \geq 1$ *and every* $u \in \mathrm{BV}(I)$.

Now we state a Theorem which connects our ergodic concepts like mixing and erodicity with the Frobenius-Perron operator. For a proof of the Theorem see [9]. Let $\mathcal{D}((X, \mathcal{A}, \mu))$ denote the space of probability density functions on the measure space $(X, \mathcal{A}, \mu)$

3.2.9 Theorem. *Let* $T : (I, \mathcal{A}, \mu) \to (I, \mathcal{A}, \mu)$ *be a measure preserving system. Let* $\mathbf{1}$ *denote the constant function equal to* 1 *everywhere. Then,*

1) T *is ergodic if and only if for any* $f \in \mathcal{D}((I, \mathcal{A}, \mu))$

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathrm{P}_T^n f \to \mathbf{1},$$

weakly in $\mathrm{L}_1(\mu)$ *as* $n \to \infty$*;*

2) T *is weakly mixing if and only if for any* $f \in \mathcal{D}((I, \mathcal{A}, \mu))$

$$\frac{1}{n}\sum_{k=0}^{n-1} |\, \mathrm{P}_T^n f - \mathbf{1}| \to 0,$$

weakly in $\mathrm{L}_1(\mu)$ *as* $n \to \infty$*;*

3) T is mixing if and only if for any $f \in \mathcal{D}((I, \mathcal{A}, \mu))$

$$\mathrm{P}_T^n f \to 1,$$

weakly in $\mathrm{L}_1(\mu)$ as $n \to \infty$;

4) T is exact if and only if for any $f \in \mathcal{D}((I, \mathcal{A}, \mu))$

$$\mathrm{P}_T^n f \to 1,$$

in the $\mathrm{L}_1(\mu)$-norm as $n \to \infty$.

Chapter 4

The two dimensional 3-rod taffy puller

In this chapter, we will use the ideas of F. D. Matthew and J-L. Thiffeault [2] and the paper of Thiffeault [1], where the two dimensional 3-rod taffy machine is represented by a linear map on the Torus to itself and by a pseudo-Anosov function on a punctured sphere. The representation of the linear map is obtained by using the braid group representation of the machine and the pseudo-Anosov representation is obtained by projecting an Anosov map to a punctured sphere.

We see the schematics of the 3-rod machine shown in Figure 4.1. Note that the first and second rods are interchanged clockwise, then the second and third are interchanged counterclockwise.

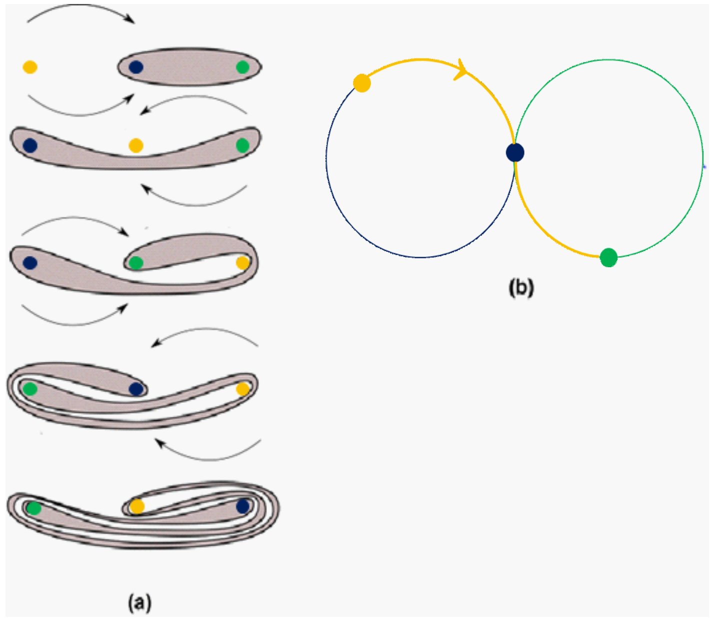


Figure 4.1: a) Motion of the taffy. b) The motion of the rods.

In [1], Thiffeault described the 3-rod taffy machine by projecting the Asonov map given in Example 2.0.17 (the cat map) to a pseudo-Anosov map on $S_{0,4}$, where $S_{0,4}$ is a topological sphere (it has genus zero) and it has 4 identified points.

$S_{0,4}$ is obtained by considering the function $\iota : \mathbb{T}^2 \to \mathbb{T}^2$, given by $\iota(x) = -x \bmod 1$, and making the quotient $\mathbb{T}^2/\iota$; *i.e.*, $S_{0,4} = \mathbb{T}^2/\iota$.

We observe that $\iota^2 = \mathrm{id}$ and has four fixed points on $\mathbb{T}^2$,

$$p_0 = (0, 0) \quad p_1 = (\tfrac{1}{2}, 0) \quad p_2 = (\tfrac{1}{2}, \tfrac{1}{2}) \quad p_3 = (0, \tfrac{1}{2}),$$

which corresponds to the 4 identified points on $S_{0,4}$.

Under the action of the Cat map f the special points (p_1, p_2, p_3) are permuted; *i.e.*, f maps (p_1, p_2, p_3) to (p_3, p_1, p_2). This points will play the role of the rods of the taffy machine. As we can see in Figure 4.2, the action of the map f after being projected to $S_{0,4}$ behaves like the 3-rod taffy machine.

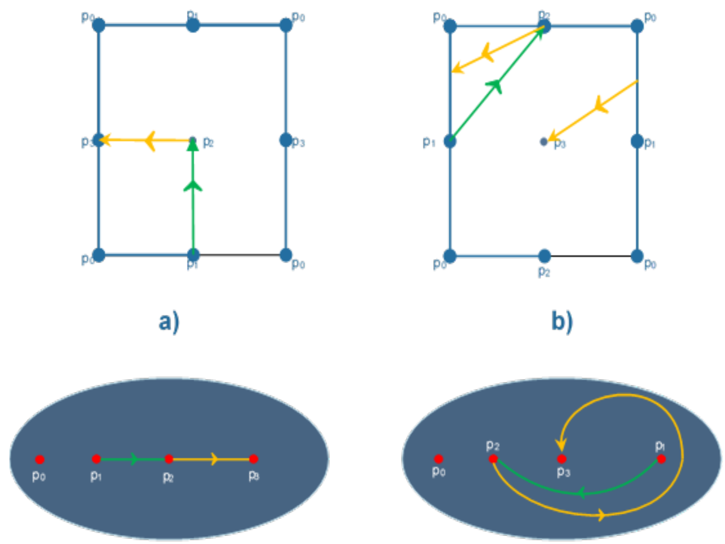


Figure 4.2: a) Two curves on $\mathbb{T}^2$, the projected curves onto $S_{0,4}$. b) The two curves transformed by the Cat map of Example2.0.17 and projected onto $S_{0,4}$.

Now we will use the ideas of F. D. Matthew and J-L. Thiffeault in [2] to represent the 3-rod machine by a linear map on $\mathbb{T}^2$ using the braid group representation.

We use the algebraic description of braids introduced in Definition 2.1.1. We order the 3 rods of the taffy machine from left to right. The generators σ_i, $i = 1, 2$ denotes the clockwise interchange of the $i-$th rod with the $i+1-$st rod along a circular path, where i is the position of a rod counting from left to right. The counter-clockwise interchange is denoted by σ_i^{-1}. A sequence of generators is an element of B_n and is called **braid word**. The Burau representation (see Definition 2.1.2) allow us to associate a 2×2 matrix to each generator. Since we are working with B_3, we have two generators and their Burau representation is given by,

$$[\sigma_1] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; [\sigma_2] = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \tag{4.1}$$

$[\sigma_1^{-1}]$ and $[\sigma_2^{-1}]$ is obtained by matrix inversion. Then we have that

$$[\sigma_1^{-1}] = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \; [\sigma_2^{-1}] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \tag{4.2}$$

We obtain the Burau representation of a braid word by multiplying the corresponding matrices together. To see in more detail the Burau representation of a braid generator see [5].

Matthew and Thiffeault in [2] used the braid word $w = \sigma_1^2\sigma_2^{-2}$ (see Figure 4.3) to represent the motion of the 3-rod machine.

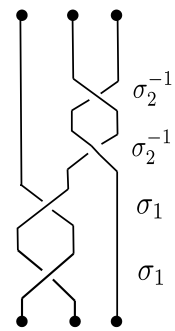


Figure 4.3: Braid group generators for three rods.

The Burau representation of the braid word w is given by,

$$[w] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}. \tag{4.3}$$

Then the linear map on the torus that describes the 3-rod taffy machine is given by,

$$g(x) = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \cdot x \bmod 1, \tag{4.4}$$

which has spectral radius $\lambda = 3 + 2\sqrt{2}$.

Chapter 5

The 4-rod taffy puller machine

In the previous chapter we introduced the standard 3-rod taffy puller machine. In this chapter and for the rest of this work we will study the standard 4-rod taffy puller machine. The model of the 4-rod machine that we work with is obtained by following the model of J. T. Halbert and J. A. Yorke presented in their paper [3].

In Figure 5.1 the schematic of the 4-rod taffy machine is shown. The machine consists of two independent arms. They are rotating around a fixed axle, and has two rods on which the taffy hangs. We will denote by S the length of the short side of the arm and by L the longer side of the arm. We can assume without loss of generality that the rescale width of the machine in resting position is 1; hence, $S, L \in [0, 1]$. Nonetheless, only S and L can be allowed to take specific values, as we will see later on.

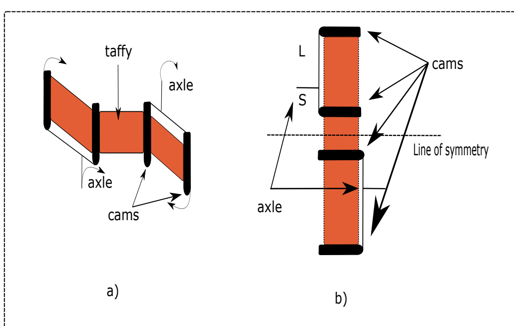


Figure 5.1: Schematic of a 4-rod taffy puller machine.

An important assumption we are going to make is that the stretching process preserves area and is linear between each pair of rods. The time is scaled so that one full revolution of the machine has period 1.

5.1 1-dimensional model

We want to preserve some topological properties of the taffy map. First, the distance between the interior rods (Γ, Figure 5.2) has to be positive. This, and the fact that the total length is 1, implies

$$L + S < \tfrac{1}{2}. \tag{5.1}$$

We will also need an overlap after the arms rotate 180° see Figure 5.2. For the taffy to be mixed as it is required we see that the overlap is necessary. To ensure this overlap we ask for $L > S + \frac{\Gamma}{2}$ which means we need

$$L > \tfrac{1}{4}. \tag{5.2}$$

Equations (5.1) and (5.2) give us the area of correct parameters (see in Figure 5.2). Since $S + L < \frac{1}{2}$ and $L > \frac{1}{4}$, we can deduce that $S < \frac{1}{4}$.

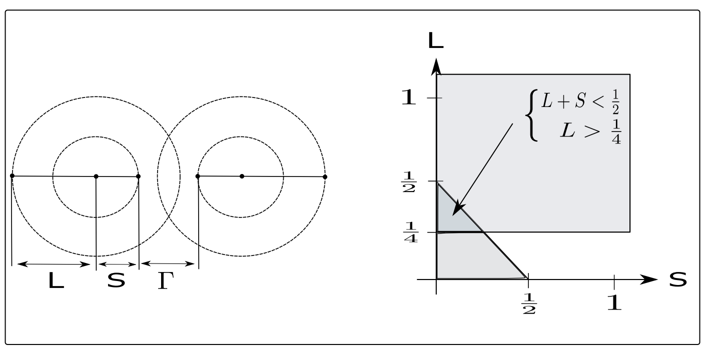


Figure 5.2: Correct values for S and L.

As we will see later on Section 4.2.2, if we change the values of the parameters S and L, we get some bad behaviour of our taffy function. In order to avoid such behaviours, we require that our parameters S and L satisfy the following inequality

$$L < \frac{1+S}{3}, \tag{5.3}$$

in the particular case when $S = 0$ that means $\frac{1}{4} < L < \frac{1}{3}$. Figure 5.3 shows the new space of good parameters.

In order to understand what is happening physically in the taffy machine, we will consider the case when $S = 0$. Then the reason why we ask L to be less than $\frac{1}{3}$, is to avoid self-intersections; in other words, we want to avoid the situation when the length of the rotating arms (L) is big enough such that the drawn circle by the arms intersects the opposite arm, like we can see in Figure 5.4 a).

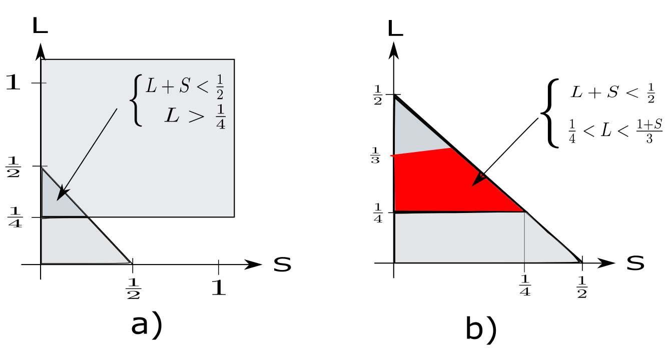


Figure 5.3: (a) the first space of parameters, (b) the new space of parameters after discarding the values for which the function T behaves badly.

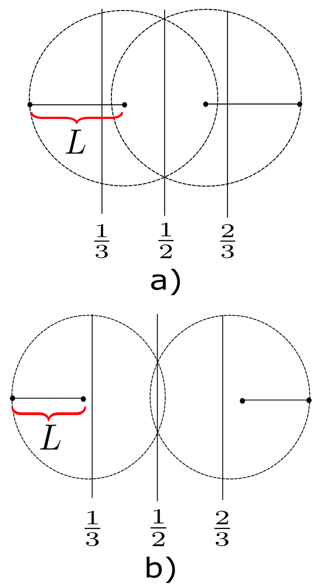


Figure 5.4: (a) Motion of the taffy machine for $S = 0$ and $L > \frac{1}{3}$, (b) Motion of the taffy machine for $S = 0$ and $L < \frac{1}{3}$.

As we can see in Figure 5.5, $c_1 = 0$, c_2 and r_1 correspond to the first cam, second cam and the axle (rotation point) of the left arm, respectively. The corresponding points of the right arm will be $c_3 = 1 - c_2$, $c_4 = 1$ and $r_2 = 1 - r_1$. $T_1(0)$ and $T_1(c_2)$ are the images under the first map T_1 of the first and second cam of the left arm, respectively. The images of the cams of the right arm are $T_1(c_3) = T_1(c_2)$ and $T_1(1) = T_1(0)$, respectively. The map T_1 shows the transition from the position where the machine is maximally extended. The map T_2, shown in Figure 5.6, begins 180° later, when the machine is minimally extended. The map $T = T_2 \circ T_1$ describes the movement of the horizontal coordinate of a point in the taffy under rotations of 180°.

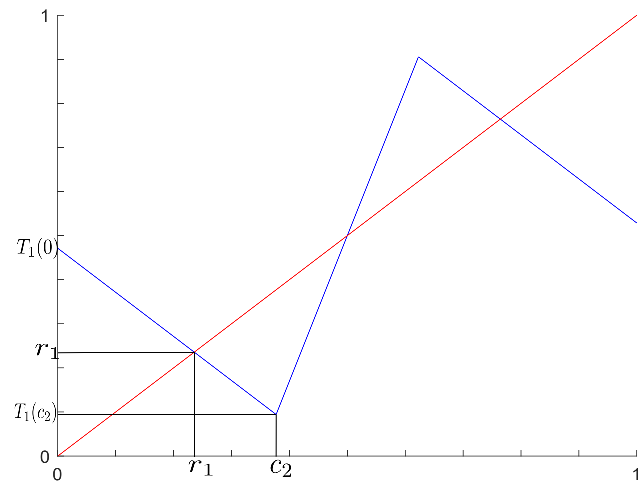


Figure 5.5: Taffy map T_1 with triangles to calculate the value of the points.

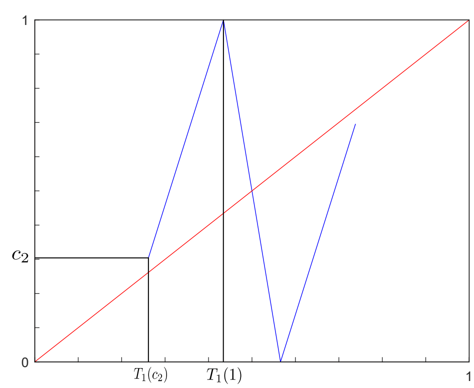


Figure 5.6: Taffy map T_2.

In order to have the equations of the maps T_1 and T_2, we have to calculate the values of c_2, r_1 and $T_1(c_2)$, since both maps are piecewise linear.

Since c_2 and r_1 are the right most cam and the axle of the left arm, respectively, and since we know that the length of the arms are $S+L$, then we can conclude that $c_2 = S+L$ and $r_1 = L$. Making the slope of the graph of T_1 in $[0, c_2]$ equal to -1, we can conclude that $T_1(c_2) = L - S$. Hence we can give an explicit formula for T_1. Using the equation of the map we can also find out that $T_1(0) = 2L$, and with this value we are ready to give an explicit formula for the map T_2. Note that we have the following order:

$$0 < T_1(c_2) < r_1 < T_1(1) < c_2 < c_3 < T_1(0) < r_2 < T_1(c_3) < 1.$$

Now that we know all the points through which the graphs of the maps T_1 and T_2 pass, we can give explicit formulas as piecewise monotone functions. Therefore the formulas of T_1 and T_2 are the following:

$$T_1(x) = \begin{cases} -x + 2L, & \text{if } x \in [0, c_2); \\ \frac{1+2S-2L}{1-2S-2L}x - \frac{2S}{1-2S-2L}, & \text{if } x \in [c_2, c_3); \\ -x + 2(1-L), & \text{if } x \in [c_3, 1], \end{cases}$$

and

$$T_2(x) = \begin{cases} \frac{1-S-L}{1+S-3L}x + \frac{2(S-SL-L^2)}{1+S-3L}, & \text{if } x \in [T_1(c_2), T_1(1)); \\ \frac{1}{1-4L}x - \frac{2L}{1-4L}, & \text{if } x \in [T_1(1), T_1(0)); \\ \frac{1-S-L}{1+S-3L}x - \frac{2(L-SL-L^2)}{1+S-3L}, & \text{if } x \in [T_1(0), T_1(c_3)]. \end{cases}$$

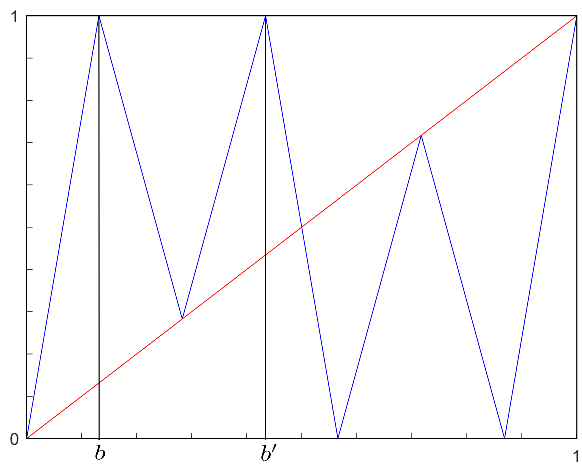


Figure 5.7: Taffy map T with points b and b'.

Now for T we need to consider the values $b = 4L-1$ and $b' = \frac{1-4L(1-L-S)}{1-2L+2S}$ which satisfy $T_1|_{[0,c_2)}(b) = T_1|_{[c_2,c_3)}(b') = T_1(1)$ (see Figure 5.7). Then the map T is given by:

$$T(x) = \begin{cases} \frac{1}{4L-1}x, & \text{if } x \in [0,b); \\ \frac{S+L-1}{S-3L+1}x + \frac{2L+2S-4L^2-4SL}{S-3L+1}, & \text{if } x \in [b,c_2); \\ \frac{1+S-3L-2S^2+2L^2}{1-S-5L-2S^2+6L^2+4SL}x + \\ \frac{4L^3-2S^2-2L^2+8SL^2+4S^2L-4SL}{1-S-5L-2S^2+6L^2+4SL}, & \text{if } x \in [c_2,b'); \\ \frac{1-2L+2S}{1-6L+8L^2+2S(4L-1)}x + \\ \frac{4L^2+4SL-2L-2S}{1-6L+8L^2+2S(4L-1)}, & \text{if } x \in [b',\frac{1}{2}); \\ 1-T(1-x), & \text{if } x \in [\frac{1}{2},1]. \end{cases}$$

Figure 5.8, 5.9, 5.10, 5.11 and 5.12 show the graphs of the maps T_1, T_2 and T for different values of S and L. As we can see in the figures, the slopes of the map T tend to be equal as S goes to 0 and L goes to $1-\frac{\sqrt{2}}{2}$.

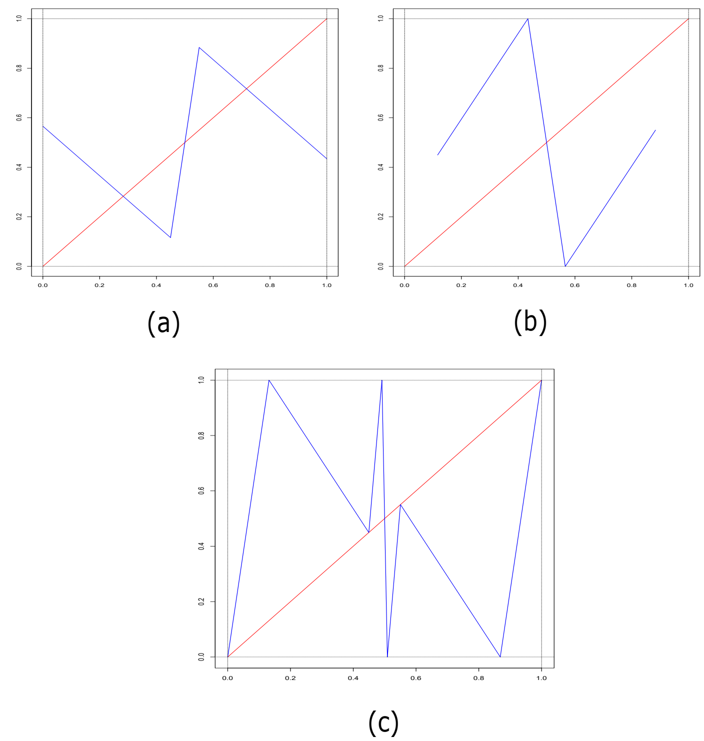
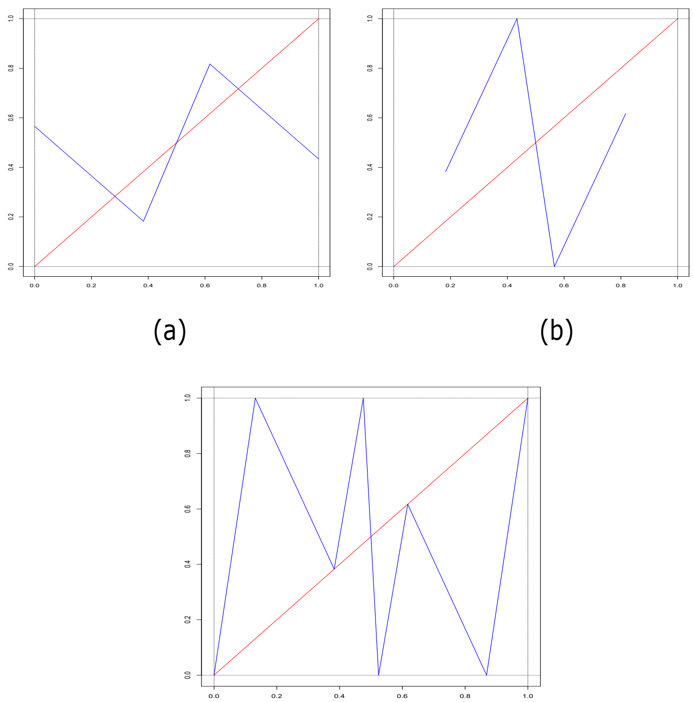


Figure 5.8: (a) The map T_1 for $S = \frac{1}{6}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$, (b) The map T_2 for $S = \frac{1}{6}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$ and (c) The map T for $S = \frac{1}{6}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$.



(a) (b)

(c)

Figure 5.9: (a) The map T_1 for $S = \frac{1}{10}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$, (b) The map T_2 for $S = \frac{1}{10}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$ and (c) The map T for $S = \frac{1}{10}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$.

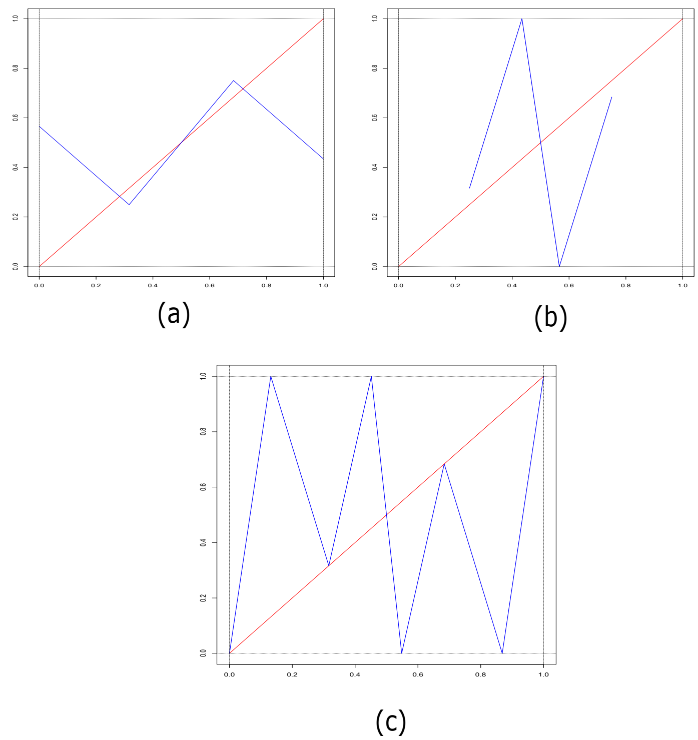


Figure 5.10: (a) The map T_1 for $S = \frac{1}{30}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$, (b) The map T_2 for $S = \frac{1}{30}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$ and (c) The map T for $S = \frac{1}{30}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$.

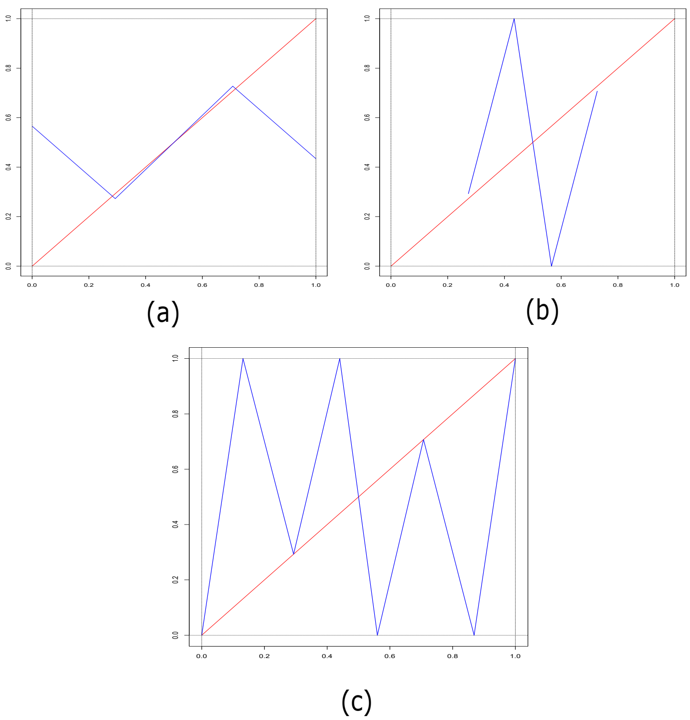


Figure 5.11: (a) The map T_1 for $S = \frac{1}{100}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$, (b) The map T_2 for $S = \frac{1}{100}$ and $L = (1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$ and (c) The map T for $S = \frac{1}{100}$ and $L = 1 - \frac{\sqrt{2}}{2}) - \frac{1}{100}$.

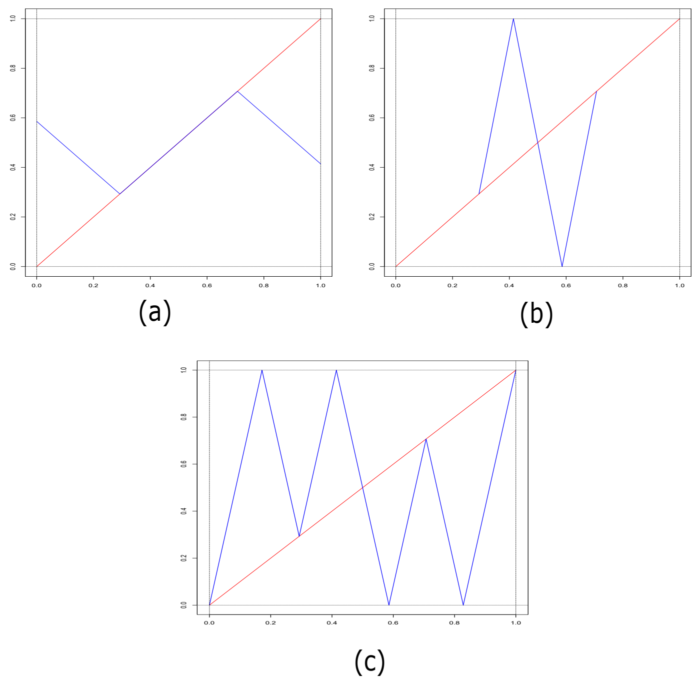


Figure 5.12: (a) The map T_1 for $S = 0$ and $L = 1 - \frac{\sqrt{2}}{2}$, (b) The map T_2 for $S = 0$ and $L = 1 - \frac{\sqrt{2}}{2}$ and (c) The map T for $S = 0$ and $L = 1 - \frac{\sqrt{2}}{2}$.

From now on we will work with the values for which the taffy map T has the same slopes, namely, $S = 0$ and $L = 1 - \frac{\sqrt{2}}{2}$. For this values of S and L, we get that $b = 3 - 2\sqrt{2}$, $c_2 = 1 - \frac{\sqrt{2}}{2}$, $b' = \sqrt{2} - 1$ and $\sigma = 3 + 2\sqrt{2}$, where σ is the slope of the map T. Hence we obtain the following equations for the map T.

$$T(x) = \begin{cases} \sigma x, & \text{if } x \in [0, b); \\ 2 - \sigma x, & \text{if } x \in [b, c_2); \\ \sigma x - \sqrt{2}, & \text{if } x \in [c_2, b'); \\ 2 + \sqrt{2} - \sigma x, & \text{if } x \in [b', \frac{1}{2}); \\ -T(1 - x), & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

To define the Markov Partition of the map T, we set $I_0 = (0, c_2)$, $I_1 = (c_2, c_3)$ and $I_2 = (c_3, 1)$ as it is shown in Figure 4.13. From the graph of the map T in Figure 5.13, we can see that $I_0 \to I_1$, $I_1 \to I_0$, $I_0 \to I_2$, $I_2 \to I_0$,

$I_1 \to I_2$ and $I_2 \to I_1$ twice each. We also have that $I_0 \to I_0$ and $I_2 \to I_2$. Last but not least, we have that $I_1 \to I_1$ thrice.

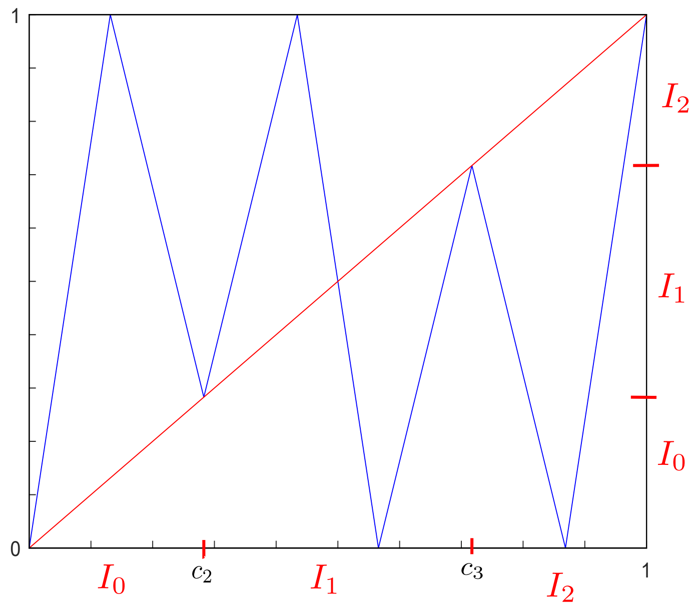


Figure 5.13: Markov partition of T.

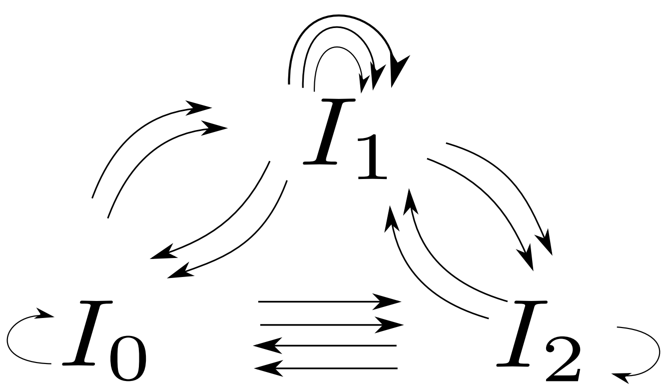


Figure 5.14: Diagram of Markov partition for T.

So in Figure 5.14 we can see the diagram that we obtain for the Markov Partition. From this we can calculate the associated transition matrix which is given by:

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

We can find out that $\det(A) = -1$ and $\mathrm{tr}(A) = 5$, then the characteristic polynomial of A is $p(\lambda) = 5\lambda^2 + 5\lambda - \lambda^3 - 1$ and the eigenvalues are $\lambda_1 = -1$, $\lambda_2 = 3 + 2\sqrt{2}$ and $\lambda_3 = 3 - 2\sqrt{2}$.

5.2 Invariant density for the taffy map T

For the invariant measure we will use the technique exposed for Yorke *et al.* [6]. We will use the transition matrix A of the Markov partition calculated in the previous section, we will multiply A by the inverse of the dilatation rate, which is the slope of the map T, and we will denote the new matrix by A'. The new matrix A' has eigenvalue 1 and the eigenvector associated to this eigenvalue corresponds to the invariant measure after normalization. Thus what we have is

$$A' = \begin{pmatrix} \frac{1}{\sigma} & \frac{2}{\sigma} & \frac{2}{\sigma} \\ \frac{2}{\sigma} & \frac{3}{\sigma} & \frac{2}{\sigma} \\ \frac{2}{\sigma} & \frac{2}{\sigma} & \frac{1}{\sigma} \end{pmatrix}$$

The eigenvalues of A' are $\lambda_1 = 1$, $\lambda_2 = 2\sqrt{2} - 3$ and $\lambda_3 = 17 - \sqrt{288}$. The corresponding eigenvectors are $e_1 = (1, \sqrt{2}, 1)$, $e_2 = (-1, 0, 1)$ and $e_3 = (1, -\sqrt{2}, 1)$, respectively. The normalization will be given by dividing e_1 by $4 - 2\sqrt{2}$. Hence we have the invariant density

$$\rho(x) = \begin{cases} \frac{1}{2} + \frac{\sqrt{2}}{4}, & \text{if } x \in [0, L) \cup [1 - L, 1]; \\ \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}), & \text{if } x \in [L, 1 - L]. \end{cases}$$

The density ρ shows the relative thickness of the taffy rescaled so that $\int_0^1 \rho(x)dx = 1$. We finally notice that the density is $\sqrt{2}$ times higher in the middle segment than on the end segments. Figure 5.15 shows the density ρ.

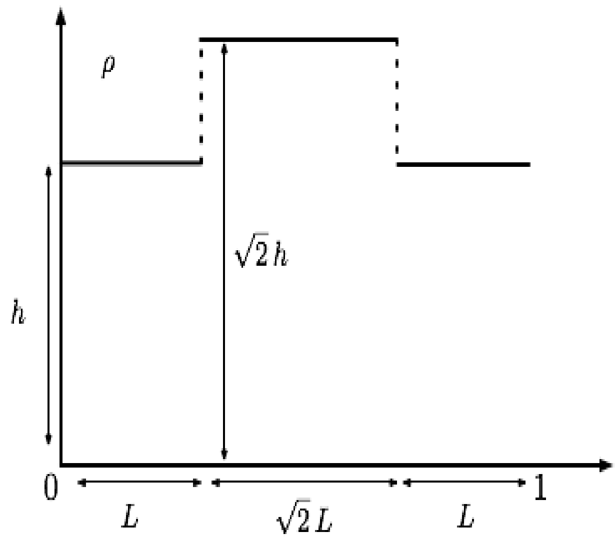


Figure 5.15: Invariant density ρ where $h = \frac{1}{2} + \frac{\sqrt{2}}{4}$ and $L = 1 - \frac{\sqrt{2}}{2}$

5.2.1 Theorem. *ρ is an invariant density*

Proof.

$$\int_0^1 \rho(x)dx = \int_0^{1-\frac{\sqrt{2}}{2}} \tfrac{1}{2} + \tfrac{\sqrt{2}}{4}dx + \int_{1-\frac{\sqrt{2}}{2}}^{\frac{\sqrt{2}}{2}} \sqrt{2}(\tfrac{1}{2} + \tfrac{\sqrt{2}}{4})dx + \int_{\frac{\sqrt{2}}{2}}^{1} \tfrac{1}{2} + \tfrac{\sqrt{2}}{4}dx.$$

but $\int_0^{1-\frac{\sqrt{2}}{2}} \frac{1}{2} + \frac{\sqrt{2}}{4}dx = \frac{1}{4}$, $\int_{1-\frac{\sqrt{2}}{2}}^{\frac{\sqrt{2}}{2}} \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4})dx = \frac{1}{2}$ and $\int_{\frac{\sqrt{2}}{2}}^{1} \frac{1}{2} + \frac{\sqrt{2}}{4}dx = \frac{1}{4}$.

So we can conclude that indeed $\int_0^1 \rho(x)dx = 1$ and ρ is a density. Now, for $A = [a, b]$ we will verify that the measure $\mu(A) = \int_a^b \rho(x)dx$ is invariant, *i.e.*, $\mu(T^{-1}A) = \mu(A)$. Set $A_1 = [\frac{a}{\sigma}, \frac{b}{\sigma})$, $A_2 = (\frac{(2+\sqrt{2})-b}{\sigma}, \frac{(2+\sqrt{2})-a}{\sigma}]$, $A_3 = [\frac{a+\sqrt{2}}{\sigma}, \frac{b+\sqrt{2}}{\sigma})$, $A_4 = (\frac{2-b}{\sigma}, \frac{2-a}{\sigma}]$, $A_5 = (\frac{(\sigma-1-\sqrt{2})-b}{\sigma}, \frac{(\sigma-1-\sqrt{2})-a}{\sigma}]$, $A_6 = [\frac{a+\sigma-1-\sqrt{2}}{\sigma}, \frac{b+\sigma-1-\sqrt{2}}{\sigma})$, $A_7 = (\frac{(\sigma-1)-b}{\sigma}, \frac{(\sigma-1)-a}{\sigma}]$ and $A_8 = [\frac{a+\sigma-1}{\sigma}, \frac{b+\sigma-1}{\sigma})$. Note that $A_1,\ A_2 \subset [0, L)$, $A_3,\ A_4,\ A_5,\ A_6 \subset [L, 1-L)$ and $A_7,\ A_8 \subset [1-L, 1]$, therefore $\mu(A_1) = \mu(A_2) = \mu(A_7) = \mu(A_8) = \frac{(b-a)}{\sigma}(\frac{1}{2} + \frac{\sqrt{2}}{4})$ and $\mu(A_3) = \mu(A_4) = \mu(A_5) = \mu(A_6) = \frac{\sqrt{2}(b-a)}{\sigma}(\frac{1}{2} + \frac{\sqrt{2}}{4})$ We will consider the following cases:

(1) $A \subseteq [0, L]$, then $\mu(A) = (b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4})$. For this case we have

$$T^{-1}A = A_1 \cup A_5 \cup A_6 \cup A_7 \cup A_8,$$

and so we get

$$\mu(T^{-1}A) = \mu(A_1) + \mu(A_5) + \mu(A_6) + \mu(A_7) + \mu(A_8) = \mu(A).$$

(2) $A \subset [1-L, 1]$, then $\mu(A) = (b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4})$. This case is analogous to the previous one, we just have to notice that $T^{-1}A = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_8$.

(3) $A \subset [L, 1-L)$, then we have that $\mu(A) = \sqrt{2}(b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4})$. For this we have

$$T^{-1}A = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_6 \cup A_7 \cup A_8,$$

thus

$$\mu(T^{-1}A) = \frac{4+3\sqrt{2}}{3+2\sqrt{2}}(b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4}) = \sqrt{2}(b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4}) = \mu(A)$$

(4) $A \subset [0, L) \cap [L, 1-L)$ or $A \subset [L, 1-L) \cap [1-L, L]$, both cases can be treated in the same fashion. We will work with the former case, for this we have $\mu(A) = (\sqrt{2}b - a + 2 - \frac{3\sqrt{2}}{2})(\frac{1}{2} + \frac{\sqrt{2}}{4})$

$$T^{-1}A = A_1 \cup A_5 \cup A_6 \cup A_7 \cup A_8 \cup \left(\tfrac{(2+\sqrt{2})-b}{\sigma}, \tfrac{(2+\sqrt{2})-L}{\sigma}\right] \cup \left[\tfrac{L+\sqrt{2}}{\sigma}, \tfrac{b+\sqrt{2}}{\sigma}\right),$$

therefore

$$\mu(T^{-1}A) = (b-a)(\frac{1}{2} + \frac{\sqrt{2}}{4}) + \int_b^L \frac{1}{2} + \frac{\sqrt{2}}{4}dx + \int_L^b \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4})dx =$$

$$(\sqrt{2}b - a + 2 - \frac{3\sqrt{2}}{2})(\frac{1}{2} + \frac{\sqrt{2}}{4}) = \mu(A).$$

We can conclude that indeed the measure $\mu(A) = \int_a^b \rho(x)dx$ is invariant. Hence the density $\rho(x)$ is an invariant density. Another way to see that $\rho(x)$ is invariant will be using the Proposition 3.0.6. First note that if $x \in [0, L)$, then $T^{-1}(x) = \{y_0, y_1, y_2, y_3, y_4\}$ with $y_0 \in I_0$, $y_1, y_3 \in I_1$ and $y_2, y_4 \in I_2$. If $x \in [1-L, 1]$, then $T^{-1}(x) = \{y_0, y_1, y_2, y_3, y_4\}$ with $y_2 \in I_2$, $y_1, y_3 \in I_1$ and

$y_0, y_4 \in I_0$. If $x \in [L, 1-L)$, then $T^{-1}(x) = \{y_0, y_1, y_2, y_3, y_4, y_5, y_6\}$ with $y_0, y_6 \in I_0$, $y_1, y_3, y_5 \in I_1$ and $y_2, y_4 \in I_2$.

Now we will consider that $x \in [0, L)$ (the case where $x \in [1-L, 1]$ is analogous). Then we have to show that

$$\hat{T}\rho(x) = \sum_{y \in T^{-1}(x)} \frac{\rho(y)}{|det(DT(y))|},$$

note that $|det(DT(y))| = 3 + 2\sqrt{2}$ for every $y \in T^{-1}(x)$, since all the slopes of T are the same. Then the above is equal to

$$\frac{1}{3+2\sqrt{2}}((\frac{1}{2} + \frac{\sqrt{2}}{4}) + 2\sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}) + 2(\frac{1}{2} + \frac{\sqrt{2}}{4})) =$$

$$\frac{1}{3+2\sqrt{2}}(3 + 2\sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4})) = \frac{1}{2} + \frac{\sqrt{2}}{4} = \rho(x).$$

Now for $x \in [L, 1-L)$, note that $(3+2\sqrt{2})(3-2\sqrt{2}) = 1$ and $(4 + 3\sqrt{2})(3-2\sqrt{2}) = \sqrt{2}$. Then if we do the corresponding calculation as the previous case we get

$$\frac{1}{3+2\sqrt{2}}(2(\frac{1}{2} + \frac{\sqrt{2}}{4}) + 3\sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}) + 2(\frac{1}{2} + \frac{\sqrt{2}}{4})) =$$

$$\frac{1}{3+2\sqrt{2}}(4 + 3\sqrt{2})(\frac{1}{2} + \frac{\sqrt{2}}{4}) = \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}) = \rho(x).$$

Therefore by Proposition 3.0.6 we can conclude that μ is invariant.

$\square$

5.2.1 Changing the partition

Now lets consider a different partition. Let $I_0 = [0, b)$, $I_1 = [b, c_2)$, $I_2 = [c_2, b')$ and $I_3 = [b', \frac{1}{2}]$ and note that $\sum_{i=0}^{3} l(I_i) = \frac{1}{2}$. Figure 5.16 shows the Markov partition for the above defined intervals. To simplify some later calculations, we will rescale and assume that $\sum_{i=0}^{3} l(I_i) = 1$. For simplicity, we just write I_i instead of $l(I_i)$.

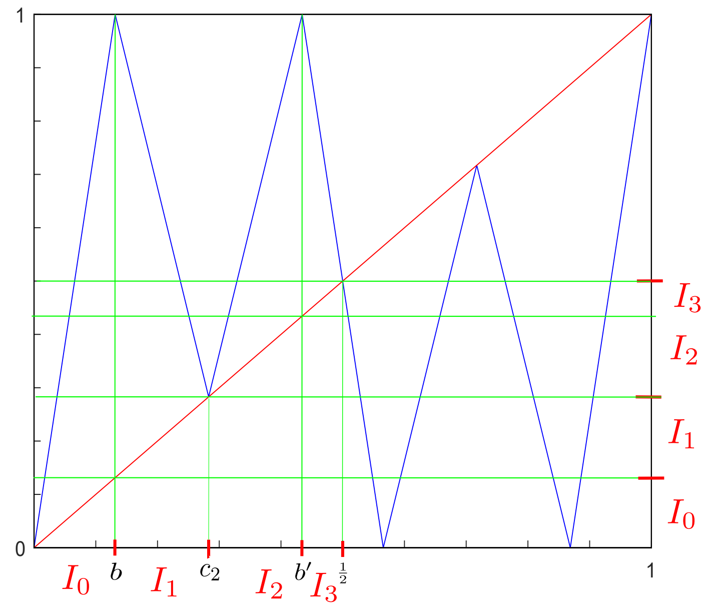


Figure 5.16: Markov partition

The associated transition matrix of this Markov partition is

$$A = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

The purpose of using this Markov partition is to consider the general case when the slopes of the taffy function are different. So we will divide the first row by the first slope of our taffy function s_0, the second row by the second slope s_2 and in the same way the third and fourth rows. Then the new matrices that we get are

$$A' = \begin{pmatrix} \frac{2}{s_0} & \frac{2}{s_0} & \frac{2}{s_0} & \frac{2}{s_0} \\ \frac{1}{s_1} & \frac{1}{s_1} & \frac{2}{s_1} & \frac{2}{s_1} \\ \frac{1}{s_2} & \frac{1}{s_2} & \frac{2}{s_2} & \frac{2}{s_2} \\ \frac{1}{s_3} & \frac{1}{s_3} & \frac{1}{s_3} & \frac{1}{s_3} \end{pmatrix}$$

Consider $w = (I_0, I_1, I_2, I_3)^\intercal$, we make the observation that the lengths of the intervals depend only in our parameters L and S. If we compute $A' \cdot w$ we obtain nice formulas for the slopes in terms of the length of the intervals, namely

$$s_0 = \frac{2}{I_0},\ s_1 = \frac{1+I_2+I_3}{I_1},\ s_2 = \frac{1+I_2+I_3}{I_2},\ s_3 = \frac{1}{I_3}. \tag{5.4}$$

Replacing the slopes in A' by the new formulas obtained in Equation 5.4, we get

$$A' = \begin{pmatrix} I_0 & I_0 & I_0 & I_0 \\ \frac{I_1}{1+I_2+I_3} & \frac{I_1}{1+I_2+I_3} & \frac{2I_1}{1+I_2+I_3} & \frac{2I_1}{1+I_2+I_3} \\ \frac{I_2}{1+I_2+I_3} & \frac{I_2}{1+I_2+I_3} & \frac{2I_2}{1+I_2+I_3} & \frac{2I_2}{1+I_2+I_3} \\ I_3 & I_3 & I_3 & I_3 \end{pmatrix}$$

We are trying to find the invariant density with respect to the transition matrix of the new Markov partition, like in the previous section, we will use the technique developed in [6] to find the invariant density. We know that 1 is an eigenvalue for the matrix A', and the left eigenvector corresponding to the eigenvalue 1 is the density after normalization.

Now, by performing $C_1 - C_2 \to C_1$, $C_4 - C_3 \to C_4$, $C_2 + aC_1 \to C_2$, $C_3 + aC_1 \to C_3$, $C_2 + dC_4 \to C_2$, $C_3 + dC_4 \to C_3$, $C_3 - 2C_2 \to C_3$ and $C_2 + C_3 \to C_2$ where the left hand side of the arrow represent the operations we made in the columns C_i, $i = 1, 2, 3, 4$, and the right hand of the arrow tells us in which column we substitute, then we reduce $A' - Id$ to the following matrix

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 + \frac{I_1}{1+I_2+I_3} & 2 - I_0 & 0 \\ 0 & \frac{I_2}{1+I_2+I_3} - 1 & -(1+I_3) & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

we note that

$$1+\frac{I_1}{1+I_2+I_3}=2-I_0 \text{ and } \frac{I_2}{1+I_2+I_3}-1=-(1+I_3). \tag{5.5}$$

The corresponding left eigenvetor is $\rho=(1+I_3,1+I_3,2-I_0,2-I_0)$. Indeed, if we compute $\rho\cdot A'$ we have

$$I_0(1+I_3)+\frac{I_1}{1+I_2+I_3}(1+I_3)+\frac{I_2}{1+I_2+I_3}(2-I_0)+I_3(2-I_0), \tag{5.6}$$

using the equalities from Equation 5.5 we get that Equation 5.6 is equal to

$$\begin{aligned}I_0(1+I_3)+\tfrac{I_1}{1+I_2+I_3}(1+I_3)+\tfrac{I_2}{1+I_2+I_3}(1+\tfrac{I_1}{1+I_2+I_3})+I_3(1+\tfrac{I_1}{1+I_2+I_3})&=\\ I_0(1+I_3)+\tfrac{I_1}{1+I_2+I_3}(1+\tfrac{I_2}{1+I_2+I_3}+2I_3)+\tfrac{I_2}{1+I_2+I_3}+I_3&=\\ I_0(1+I_3)+\tfrac{I_1}{1+I_2+I_3}(1+I_3)&=\\ (I_0+\tfrac{I_1}{1+I_2+I_3})(1+I_3)&=\\ 1+I_3,&\end{aligned}$$

for the first two columns of A', which correspond to the first two entries of the vector ρ. For the remaining two columns, which correspond to the last two entries of ρ, we get

$$I_0(1+I_3)+\frac{2I_1}{1+I_2+I_3}(1+I_3)+\frac{2I_2}{1+I_2+I_3}(2-I_0)+I_3(2-I_0), \tag{5.7}$$

again using the equalities from Equation 5.5 we get that Equation 5.7 is equal to

$$\begin{aligned}\tfrac{2I_2}{1+I_2+I_3}(2-I_0)+I_3(2-I_0)+I_0(1-\tfrac{I_2}{1+I_2+I_3})+\tfrac{2I_1}{1+I_2+I_3}(1-\tfrac{I_2}{1+I_2+I_3})&=\\ I_0+\tfrac{2I_1}{1+I_2+I_3}-\tfrac{I_0I_2}{1+I_2+I_3}+\tfrac{2I_2}{1+I_2+I_3}(2-I_0-\tfrac{I_1}{1+I_2+I_3})+I_3(2-I_0)&=\\ (2-I_0)+\tfrac{I_2}{1+I_2+I_3}(2-I_0)+I_3(2-I_0)&=\\ (2-I_0)(1+I_3+\tfrac{I_2}{1+I_2+I_3})&=\\ 2-I_0.&\end{aligned}$$

Let $u = \rho \cdot w = (1 + I_3)(I_0 + I_1) + (2 - I_0)(I_2 + I_3)$. Then, the density will be $\rho = (\frac{1+I_3}{u}, \frac{1+I_3}{u}, \frac{2-I_0}{u}, \frac{2-I_0}{u})$; *i.e.*,

$$\rho(x) = \begin{cases} \frac{1+I_3}{u}, & \text{if } x \in I_0 \cup I_1; \\ \frac{2-I_0}{u}, & \text{if } x \in I_2 \cup I_3. \end{cases} \tag{5.8}$$

5.2.2 Theorem. *$\rho(x)$ is an invariant density.*

Proof. We will first see that $\int_0^1 \rho(x)dx = 1$. Indeed, we have that

$$\int_0^1 \rho(x)dx = I_0(\frac{1+I_3}{u}) + I_1(\frac{1+I_3}{u}) + I_2(\frac{2-I_0}{u}) + I_3(\frac{2-I_0}{u}) = \frac{1}{u}((1 + I_3)(I_0 + I_1) + (2 - I_0)(I_2 + I_3)) = 1.$$

Therefore, $\rho(x)$ is a density. Now we will make use of Proposition 3.0.6 to prove that $\rho(x)$ is an invariant density. First notice that for any $x \in I_0 \cup I_1$, then $T^{-1}(x) = \{y_0, y_1, y_2, y_3, y_4\}$ with $y_0, y_4 \in I_0$, $y_1 \in I_1$, $y_2 \in I_2$ and $y_3 \in I_3$. If $x \in I_2 \cup I_3$, then $T^{-1}(x) = \{y_0, y_1, y_2, y_3, y_4, y_5, y_6\}$ with $y_0, y_6 \in I_0$, $y_1, y_5 \in I_1$, $y_2, y_4 \in I_2$ and $y_3 \in I_3$. From Equation (2.1) and the above observation, for $x \in I_0 \cup I_1$ we get

$$\hat{T}\rho(x) = \frac{2I_0}{2}\frac{1+I_3}{u} + \frac{I_1}{1+I_2+I_3}\frac{1+I_3}{u} + \frac{I_2}{1+I_2+I_3}\frac{2-I_0}{u} + I_3\frac{2-I_0}{u} = \frac{1}{u}((1+I_3)(I_0 + \frac{I_1}{1+I_2+I_3}) + (2-I_0)(I_3 + \frac{I_2}{1+I_2+I_3})) = \frac{1+I_3}{u},$$

the last equality above is just the calculation we did for Equation 5.6. In the same way for $x \in I_2 \cup I_3$ and using the calculation we did for Equation 5.7 we get

$$\hat{T}\rho(x) = \frac{2I_0}{2}\frac{1+I_3}{u} + \frac{2I_1}{1+I_2+I_3}\frac{1+I_3}{u} + \frac{2I_2}{1+I_2+I_3}\frac{2-I_0}{u} + I_3\frac{2-I_0}{u} = \frac{1}{u}((1+I_3)(I_0 + \frac{2I_1}{1+I_2+I_3}) + (2-I_0)(I_3 + \frac{2I_2}{1+I_2+I_3})) = \frac{2-I_0}{u}.$$

Therefore, $\rho(x)$ is invariant. $\square$

5.2.3 Remark. *With the work we just did now, given any good value for our parameters S and L we can calculate the invariant density of our taffy machine.*

Let's check what we just did. If we consider the case where all slopes are the same, namely $\sigma = 3+2\sqrt{2}$, we find out that $I_0 = 3-2\sqrt{2}$, $I_1 = \frac{3}{2}\sqrt{2}-2$, $I_2 = \frac{3}{2}\sqrt{2}-2$, $I_3 = \frac{3}{2}-\sqrt{2}$, we have that $\sum_{i=0}^{3} I_i = \frac{1}{2}$, so we rescale by multiplying by two each I_i. So we get that $I_0 = 6-4\sqrt{2}$, $I_1 = 3\sqrt{2}-4$, $I_2 = 3\sqrt{2}-4$ and $I_3 = 3-2\sqrt{2}$. Lets consider the vector w whose entries are the lengths of the intervals of our partition after rescaling; *i.e.*, $w = (I_0, I_1, I_2, I_3)^{\intercal}$. According to what we just did

$$u = (4-2\sqrt{2})(2-\sqrt{2}) + (4\sqrt{2}-4)(\sqrt{2}-1) = 8(3-2\sqrt{2}),$$

$$\frac{1+I_3}{u} = \frac{4-2\sqrt{2}}{8(3-2\sqrt{2})} = \frac{1}{2} + \frac{\sqrt{2}}{4}$$

and

$$\frac{2+I_0}{u} = \frac{4\sqrt{2}-4}{8(3-2\sqrt{2})} = \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}).$$

Hence, the invariant density $\rho(x)$ is given by

$$\rho(x) = \begin{cases} \frac{1}{2} + \frac{\sqrt{2}}{4}, & \text{if } x \in I_0 \cup I_1; \\ \sqrt{2}(\frac{1}{2} + \frac{\sqrt{2}}{4}), & \text{if } x \in I_2 \cup I_3. \end{cases}$$

5.2.2 Changing the slopes

For this part we will work with different values for the parameters S and L which will produce that our taffy function T has different slopes in each of their pieces of monotonicity. Let's consider $S = \frac{1}{10}$ and $L = \frac{1}{3}$. Then the taffy function T is given by

$$T(x) = \begin{cases} 3x, & \text{if } x \in [0, b); \\ -\frac{17}{3}x + \frac{130}{9}, & \text{if } x \in [b, c_2); \\ \frac{68}{3}x - \frac{169}{18}, & \text{if } x \in [c_2, b'); \\ -12x + \frac{13}{2}, & \text{if } x \in [b', \frac{1}{2}); \\ 1 - T(1-x), & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

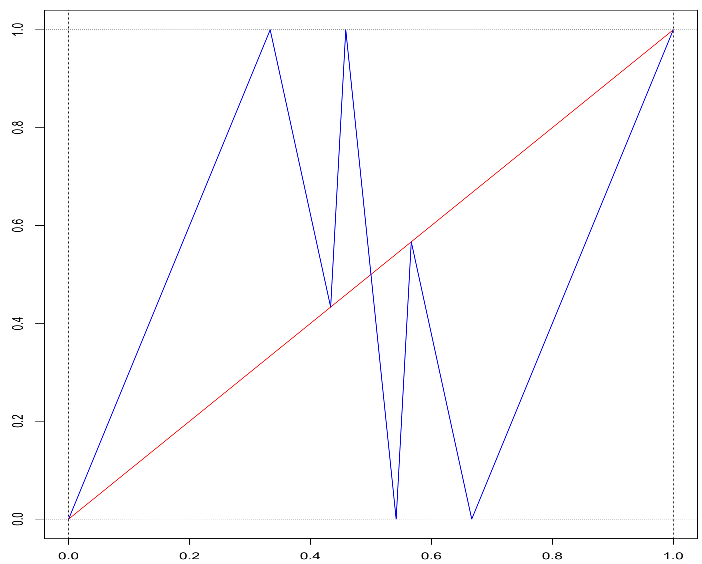


Figure 5.17: Graph of the taffy function for $S = \frac{1}{10}$ and $L = \frac{1}{3}$

where $b = \frac{1}{3}$, $c_2 = \frac{13}{30}$ and $b' = \frac{11}{24}$. Let $s_0 = s_7 = 3$, $s_1 = s_6 = \frac{17}{3}$, $s_3 = s_5 = \frac{68}{3}$ and $s_4 = 12$ be the absolute value of the slopes of each of the 7 pieces of monotonicity of the function T, in Figure 5.17 we can see the graph of the taffy function. Like in the previous subsection, we will consider the partition $I_0 = [0, b)$, $I_1 = [b, c_2)$, $I_2 = [c_2, b')$ and $I_3 = [b', \frac{1}{2}]$. Note that $I_0 = \frac{1}{3}$, $I_1 = \frac{1}{10}$, $I_2 = \frac{1}{40}$, $I_3 = \frac{1}{24}$ and $\sum_{i=0}^{3} l(I_i) = \frac{1}{2}$. We rescale again by multiplying by two each I_i, we get $I_0 = \frac{2}{3}$, $I_1 = \frac{1}{5}$, $I_2 = \frac{1}{20}$, $I_3 = \frac{1}{12}$ also we have that $w = (\frac{2}{3}, \frac{1}{5}, \frac{1}{20}, \frac{1}{12})^\intercal$ and $u = \frac{67}{60}$. So the invariant density for this case is given by

$$\rho(x) = \begin{cases} \frac{65}{67}, & \text{if } x \in I_0 \cup I_1; \\ \frac{80}{67}, & \text{if } x \in I_2 \cup I_3. \end{cases}$$

As we mentioned at the beginning of this chapter, we can find bad behaviour in the taffy function when the parameters do not fulfil the Equation 5.3. As an example of this bad behaviour if we consider $S = \frac{1}{100}$ and $L = \frac{2}{5}$, the taffy map is given by

$$T(x) = \begin{cases} \frac{5}{3}x, & \text{if } x \in [0, b); \\ \frac{59}{19}x - \frac{82}{95}, & \text{if } x \in [b, c_2); \\ -\frac{649}{171}x - \frac{1{,}681}{855}, & \text{if } x \in [c_2, b'); \\ -\frac{55}{27}x + \frac{41}{27}, & \text{if } x \in [b', \frac{1}{2}); \\ 1 - T(1-x), & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

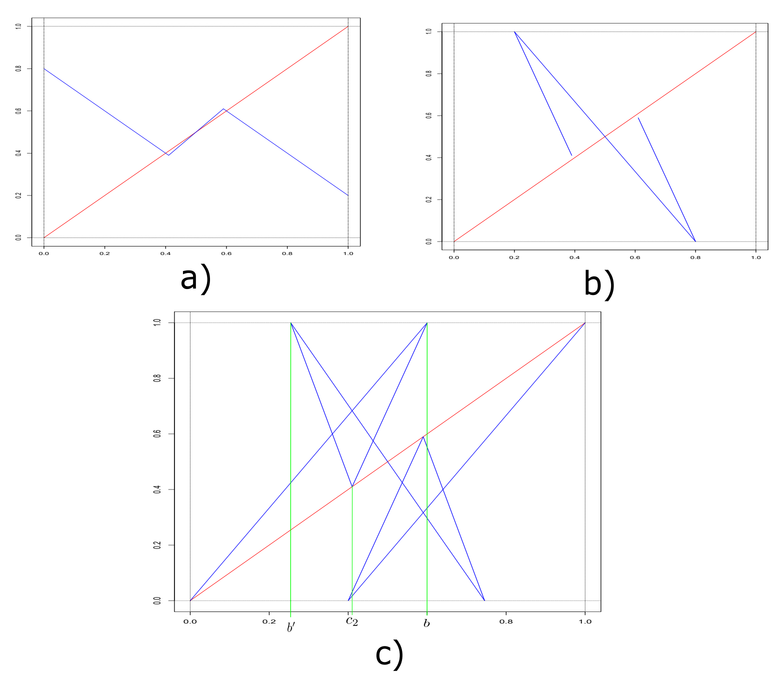


Figure 5.18: (a) The map T_1 for $S = \frac{1}{100}$ and $L = \frac{2}{5}$, (b) The map T_2 for $S = \frac{1}{100}$ and $L = \frac{2}{5}$ and (c) The map T for $S = \frac{1}{100}$ and $L = \frac{2}{5}$.

Figure 5.18 shows the graph of the taffy function, we can observe that the map changes the order of the points b and b' and the sign of the slopes. The partition $I_0 = [0, b)$, $I_1 = [b, c_2)$, $I_2 = [c_2, b')$ and $I_3 = [b', \frac{1}{2}]$ does not work any more for the calculation of the invariant density.

We will prove the last result of this chapter, namely, the mixing property of the taffy function.

5.2.4 Theorem. *The taffy function T is exponentially mixing*

Proof. To prove that T is mixing we will make use of the Proposition 3.0.17. Let $f \in \mathrm{L}_1(\mu)$ and $g \in \mathrm{L}_\infty(\mu)$, where $\mu(A) = \int_A \rho(x)dx$, we will show that

$$\int_I f \cdot g \circ T^n d\mu - \int_I f d\mu \int_I g d\mu \to 0, \tag{5.9}$$

using invariance of the measure μ, we have that Equation (5.9) is equivalent to

$$\int_I f \cdot g \circ T^n d\mu - \int_i \left(\int_I f d\mu \right) \cdot g \circ T^n d\mu = \int_I \left(f - \int_I f d\mu \right) \cdot g \circ T^n d\mu, \tag{5.10}$$

now let $v = f - \int_I f d\mu$ and using property 5 of Theorem 3.2.2, we have that Equation (5.10) is equivalent to

$$\int_I (\mathrm{P}_T^n v) \cdot w d\mu.$$

Now using Hölder inequality (Theorem 3.0.10), we see that

$$\int_I (\mathrm{P}_T^n v) \cdot w d\mu \leq \|\mathrm{P}_T^n v\|_1 \|w\|_\infty .$$

We notice that $\mathrm{Q}\, v = 0$ since $\mathrm{P}_T\, v = \int_I f d\mu$, where Q is as in Remark 3.2.8. Using the above observation and the decomposition of the operator P_T that we have in the Remark 3.2.8, we get the following

$$\begin{aligned} \|\mathrm{P}_T^n v\|_1 \|w\|_\infty = \|\mathrm{P}_T^n v - \mathrm{Q}\, v\|_1 \|w\|_\infty = \|\mathrm{S}^n v\|_1 \|w\|_\infty \leq \|\mathrm{S}\|_{\mathrm{BV}}^n \|v\|_1 \|w\|_\infty \\ \leq \eta^n \|v\|_1 \|w\|_\infty , \end{aligned}$$

where $\eta \in (0, 1)$. What we have then is the following

$$\int_I (\mathrm{p}_T^n v) \cdot w d\mu \leq \eta^n \|v\|_1 \|w\|_\infty .$$

Therefore, we can conclude that $\int_I f \cdot g \circ T^n d\mu - \int_I f d\mu \int_I g d\mu \to 0$. Hence, the map T is exponentially mixing. $\square$

Chapter 6

2 dimensional map

In this section we describe the model of the taffy machine in two dimensions as the natural extension of our 1-dimensional taffy map T and as a linear map on the Tourus. First, we will describe the space on the plane where the 2-dimensional map acts, this construction is taken from the work of J. T. Halbert and J. A. Yorke [3]. As shown in Figure 6.1, we cut the taffy at each rod and eliminate the holes that are produced by the rods. The cut edges are put back together. Next, we identify the points along the cut edges and rescale the vertical coordinate so that the height is the same as the width. The cuts are vertical edges in the taffy and in the limit we get the lower right picture. The points located at the center of the four vertical edges are the four rods of the machine.

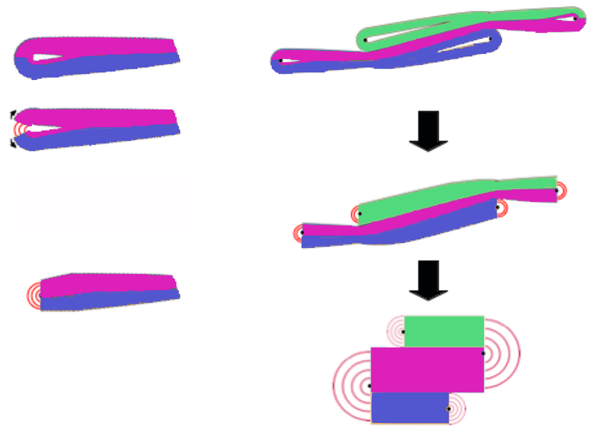


Figure 6.1: Creating the Taffy block map. Each rod is treated as shown on the left.

The proportions of the block in Figure 6.1 can be shown in Figure 6.2. The dynamics behave according to the 1-dimensional map T in the x coordinate. The points located near a rod do not separate from the rod under the action of the map map; hence, in the limit the rods become fixed points. The described process produces the action of the map shown in Figure 6.4. The thickness of the taffy and the invariant density ρ are proportional to each other. We will denote by X the subset of the plane shown in Figure 6.2. We construct the topological space $\tilde{X}$ by making the identifications (red semi-circles in Figure 6.3) of the sides of X.

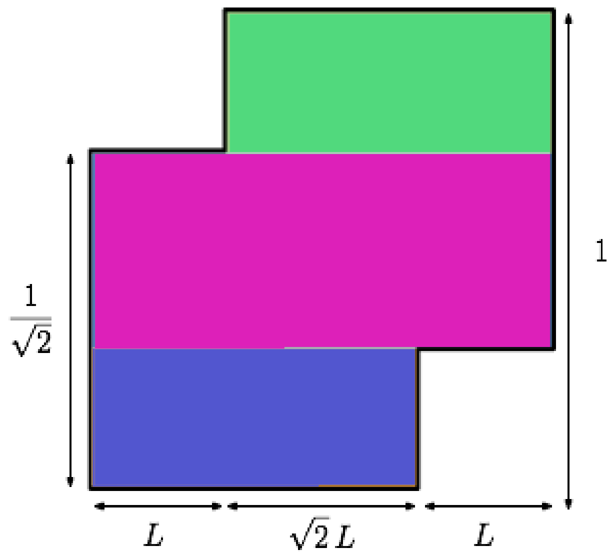


Figure 6.2: The set X, vertically rescaled taffy, where $L = 1 - \frac{\sqrt{2}}{2}$.

Before we continue with the study of the 2 dimensional taffy map we will give a description of the invariant domain. The continuity of the taffy is preserved by identifying the vertical segments above and below each external fixed point (f_1, f_2, f_4 and f_5 in Figure 6.4). We can see this identification in Figure 6.3. We make the observation that the two-dimensional map is one-to-one except on the edges where the map is two-to-one.

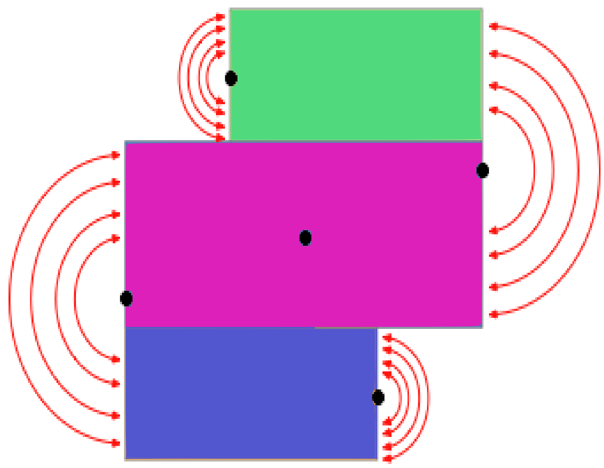


Figure 6.3: The semi-circles connect points that are identified with each other. The black dots represent fixed points. We write $\tilde{X}$ for this set after making the identifications shown here.

Let F denote the taffy map shown in Figure 6.4. It acts piecewise linearly on each coloured region. As we already mentioned the points f_1, f_2, f_3, f_4 and f_5 are the fixed points of the map F. We know that in the x-coordinate the map F behaves according to our 1 dimensional taffy map T, the fixed points of T are 0, L, $\frac{1}{2}$, $1-L$ and 1. We also know that the fixed points of F correspond to the cams of the taffy machine. Therefore, the height of the points (the y-coordinate), by the construction of the block, are located at the center of the four vertical edges. Hence we have that $f_1 = (L, 1-\frac{L}{2})$, $f_2 = (1, 1-\frac{1}{2\sqrt{2}})$, $f_3 = (\frac{1}{2}, \frac{1}{2})$, $f_4 = (0, \frac{1}{2\sqrt{2}})$ and $f_5 = (\frac{1}{\sqrt{2}}, \frac{L}{2})$. Since F is piecewise linear in each of the regions we can give an explicit formula for F; namely,

$$F(x,y) = \begin{cases} (T(x), \sigma^{-1}(\frac{L}{2}(\sigma+1) - y)), & \text{if } y \in [0, \frac{L}{2}]; \\ (T(x), \sigma^{-1}(\frac{1}{2\sqrt{2}}(\sigma+1) - y)), & \text{if } y \in (\frac{L}{2}, \frac{1}{2\sqrt{2}}]; \\ (T(x), \sigma^{-1}(\frac{\sigma-1}{2} + y)), & \text{if } y \in (\frac{1}{2\sqrt{2}}, \frac{1}{2}]; \\ (T(x), \sigma^{-1}(1 - \frac{1}{2\sqrt{2}}(\sigma+1) - y)), & \text{if } y \in (\frac{1}{2}, 1 - \frac{1}{2\sqrt{2}}]; \\ (T(x), \sigma^{-1}(1 - \frac{L}{2}(\sigma+1) - y)), & \text{if } y \in (1 - \frac{1}{2\sqrt{2}}, 1]. \end{cases}$$

It has derivative

$$DF = \pm \begin{pmatrix} \sigma & 0 \\ 0 & \sigma^{-1} \end{pmatrix}, \tag{6.1}$$

everywhere except at the external fixed points, where $\sigma = 3 + 2\sqrt{2}$. Since we have rotations of 180° in some regions an in others not, we need the plus and minus sign. Heuristically F stretches the green region horizontally away from f_1, stretching here means following the identifications.

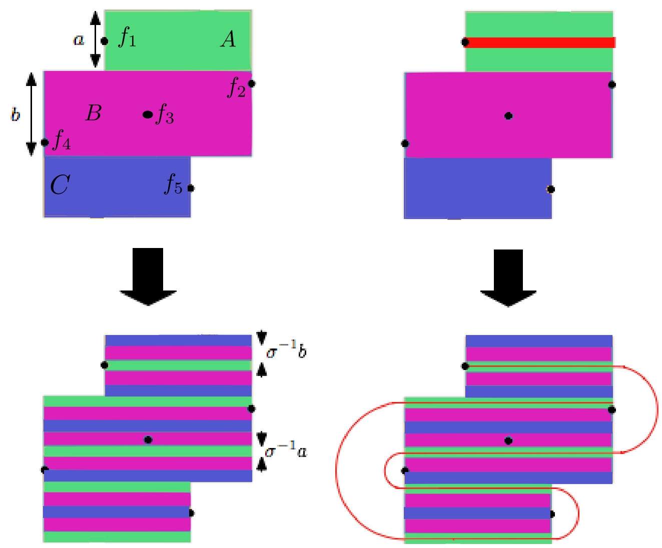


Figure 6.4: The action of the 2 dimensional map. The coloured strips in the lower half of the figure are thinner by a factor of σ.

We show in Figure 6.5 the transition graph of the 2-dimensional map. It shows how many times each rectangle stretches across the others. We note that this partitions is actually a Markov partition.

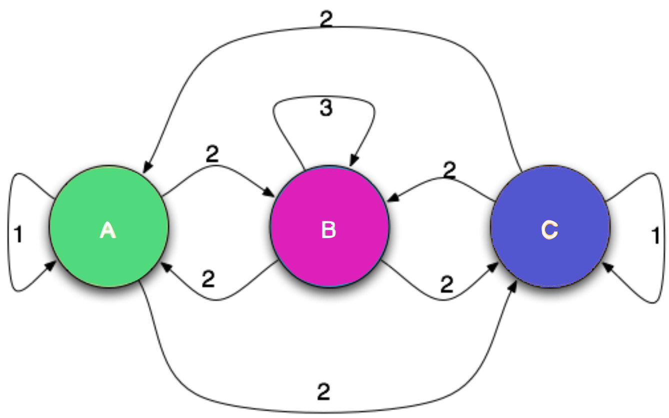


Figure 6.5: Transition graph for 2 dimensional taffy map. The numbers with the bent arrows represent how many times the image of a region stretches across the region the arrow points to.

6.0.1 Observation. *The 2 dimensional taffy map described above just corresponds to the case where all the slopes of the 1 dimensional map are the same; i.e., the case where* $S = 0$ *and* $L = 1 - \frac{\sqrt{2}}{2}$ *(see Section 1 of chapter 4). It remains to work the general case of the 2 dimensional taffy map, a work that can be done in the future.*

6.1 The 2 dimensional map as the natural extension

Let us recall the definition of natural extension (Definition 3.0.8). A system $(X, \mathcal{A}, \lambda, F)$ is called a **natural extension** of a measure preserving system $(Y, \mathcal{B}, \mu, T)$ if there are $X' \in \mathcal{A}$, $Y' \in \mathcal{B}$ and $\phi : X' \to Y'$ such that

1.- $\lambda(X') = 1$, $\mu(Y') = 1$;

2.- $\phi : X' \to Y'$ is a measurable surjection;

3.- ϕ is measure preserving; *i.e.*, $\lambda(\phi^{-1}(B)) = \mu(B)$ for every $B \in \mathcal{B}$;

4.- $\phi \circ F = T \circ \phi$;

5.- $F : X' \to X'$ is invertible

6.1.1 Example. *The baker transformation is given by*

$$B(x,y) = \begin{cases} (2x, \frac{y}{2}), & \text{if } x \in [0, \frac{1}{2}); \\ \\ (2-2x, \frac{1-y}{2}), & \text{if } x \in [\frac{1}{2}, 1), \end{cases}$$

being the natural extension of the doubling map (Example 3.2.3).

The two dimensional taffy map $(X, \mathcal{A}, \lambda, F)$ is the natural extension of the 1 dimensional taffy map $([0,1], \mathcal{B}, \mu, T)$, where X is the set described in Figure 6.2, $\mathcal{A}$ is the Borel σ-algebra of the unit square restricted to X, λ is the Lebesgue measure in the unit square restricted to X and normalized such that $\lambda(X) = 1$, $\mathcal{B}$ is the Borel σ-algebra of the unit interval and μ is the measure with density ρ described in Theorem 5.2.1. The sets of full measure that we will use are the whole spaces;*i.e.*, $X' = X$ and $Y = Y'$ and the function ϕ is the projection $\pi : X \to [0,1]$. So we have that the following diagram commute;

$$\begin{tikzcd} X \arrow[r, "F"] \arrow[d, "\pi"'] & X \arrow[d, "\pi"] \\ {[0,1]} \arrow[r, "T"] & {[0,1]} \end{tikzcd}$$

Therefore, all points 1 to 5 described above are satisfied. Hence, the 2 dimensional taffy map F is the natural extension of the 1 dimensional taffy map T.

6.2 Description as a linear map on the torus

For the 2 dimensional case, we will describe it with a linear map on the torus as it was done with the 3-rod taffy model. We will give the braid group representation of the 4-rod machine.

Let's think about the 4-rod taffy machine now. Figure 6.1 shows the rod motion of the machine. A full period of the puller requires all the rods to return to their initial position.

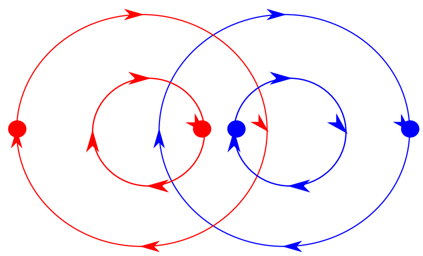


Figure 6.6: Rod motion

As we did in Chapter 4, we will use the algebraic representation of braids. For this case we are going to work with B_4. Then we have three generators, namely, σ_1, σ_2 and σ_3 and their inverses. Using the Burau representation again (see Definition 2.1.2) we associate a 3×3 matrix to each generator. Hence we have the following,

$$[\sigma_1] = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ [\sigma_2] = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \ [\sigma_3] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}. \quad (6.2)$$

For the 4-rod taffy machine the braid word that represent the action of the machine is $w = \sigma_1\sigma_3\sigma_2^2\sigma_1\sigma_3$ see Figure

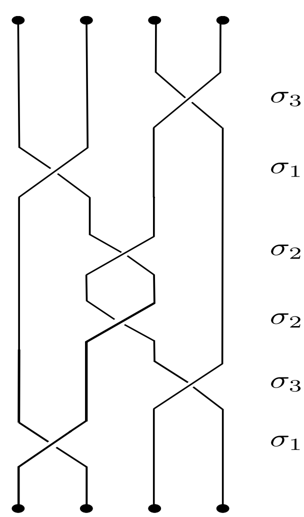


Figure 6.7: The braid representation of the 4-rod machine.

As we performed in Chapter 4, we have that the Burau representation of the braid word w is given by multiplying the the matrices associated to the generators. Therefore we have that;

$$[w] = \begin{pmatrix} -1 & -2 & 2 \\ -2 & -3 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \tag{6.3}$$

its has eigenvalues $\lambda_1 = 1$, $\lambda_2 = -3+2\sqrt{2}$ and $\lambda_3 = -3-2\sqrt{2}$ with associated eigenvectors $(1,0,1)$, $(-1,\sqrt{2},1)$ and $(-1,-\sqrt{2},1)$, respectively.

Now we want to reduce this 3×3 matrix to a 2×2 matrix. We first note that since $[w]$ is symmetric, its eigenspaces are mutually perpendicular. Let's call E_1 to the eigenspace corresponding to the eigenvalue $\lambda = 1$. Then $E_1 = \operatorname{span}\{(1,0,1)^\intercal\}$, observe that $\dim E_1 = 1$. Thus we will restrict our matrix $[w]$ to the orthogonal complement of E_1; *i.e.*, to $E_1^\perp = \operatorname{span}\{(1,1,-1)^\intercal\} = \operatorname{span}\{(1,0,-1)^\intercal, (0,1,0)^\intercal\}$, note that $\dim E_1^\perp = 2$.

Using the basis $\{(1,0,-1)^\intercal, (0,1,0)^\intercal\}$, we have that the restriction of $[w]$ to $E_1^\perp$, denoted by $[w]|_{E_1^\perp}$, is given by

$$[w]|_{E_1^\perp} = \begin{pmatrix} -3 & -4 \\ -2 & -3 \end{pmatrix}, \tag{6.4}$$

which has eigenvalues $\lambda = 3 \pm 2\sqrt{2}$, the desired eigenvalues.

We can conclude that the 4-rod taffy machine is represented by the linear map on the torus

$$h(x) = \begin{pmatrix} -3 & -4 \\ -2 & -3 \end{pmatrix} \cdot x \bmod 1. \tag{6.5}$$

6.2.1 Observation. *We conclude this work with the observation that the 4-rod machine and the 3-rod machine are conjugated. We have the following conjugacies:*

$$\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \tag{6.6}$$

and

$$\begin{pmatrix} -3 & -4 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{6.7}$$

Therefore we have that;

$$\begin{pmatrix} -3 & -4 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}. \tag{6.8}$$

Bibliography

[1] Thiffeault, Jean-Luc, *The mathematics of taffy pullers*, The Mathematical Intelligencer, 40, 26-35, 2018.

[2] Finn, D. Matthew and Thiffeault, Jean-Luc, *Topological Optimization of Rod-Stirring Devices*, SIAM Review, 53, 723-743, 2011.

[3] Halbert, James T. and Yorke, James A., *Modeling a chaotic machine's dynamics as a linear map on a "square sphere"*, Topology Proceedings, 2014.

[4] Fathi, Albert and Laudenbach, François and Poénaru, Valentin, *Thurston's Work on Surfaces (MN-48)*, Princeton University Press, Vol. 48, 2012.

[5] Kolev, Boris, *Topological entropy and Burau representation*, arXiv preprint math/0304105, 2003.

[6] Alligood, Kathleen, D. Sauer, Tim and A. Yorke, James, *Chaos: An Introduction to Dynamical Systems*, Physics Today, 50, 1997.

[7] Birman, Joan S, *Braids, links and mapping class groups*, Bull. Amer. Math. Soc, 82, 1976.

[8] Tao, Terence, *An introduction to measure theory*, American Mathematical Society Providence, Vol. 126, 2011.

[9] Boyarsky, Abraham and Gora, Pawel *Laws of chaos: invariant measures and dynamical systems in one dimension*, Springer Science & Business Media 2012.

[10] Natanson, I.P., *Theory of functions of a real variable*, Frederick Ungar Punl. Co., 1955.

[11] Walters, Peter, *An introduction to ergodic theory*, Springer Science & Business Media, Vol. 79 2000.

[12] Schaefer, H.H., *Banach lattices and positive operators*, Die Grundlehren der mathematischen Wiessenschaften, Springer-Verlag, New York-Heidelberg 1974.

[13] Macías, Sergio, *Topics on Continua*, Pure and Applied Mathematics Series, Vol. 82, Chapman & Haul/CRC, Taylor & Francis Group, Boca Raton, London, New York, Singapore, 2005.

[14] Genovese, Peter, *Chew on this: 125 years later, Jersey shore still daffy over salt water taffy*, The Newark Star Ledger, 2013.

[15] Shenkman, Ken, *History of Salt Water Taffy*, www.bulkcandystore.com, 2012.

Index