# On Variations of the SRB Entropy of the Expanding Map on the Circle

Richard (Rick) G. Freedman
Mentored by Dr. Miaohua Jiang

May 24, 2012

**Abstract**

Many people are familiar with the geometrical shape called the circle. Based on this figure, the circle space $\mathbb{S}^1$ connects the endpoints of the interval $[0, 1]$ together so that $0 \equiv 1 \pmod 1$. On this space, the expanding map $f : \mathbb{S}^1 \to \mathbb{S}^1$ stretches an initial distribution of points along the circle and then rewraps the lengthened distribution tightly about the circle; overlapping regions are compressed together to yield the new distribution along the circle. By iteratively performing the expanding map, we get a discrete dynamical system whose orbits are chaotic. The entropy is an observation of the complexity of this chaos with respect to a given measure. We study variations of the entropy of expanding maps that are small perturbations of the uniformly expanding map on $\mathbb{S}^1$ with respect to the physical measure, also known as the Sinai Ruelle Bowen (SRB) invariant measure. Due to the complexity of these computations, we discuss methods for computing this entropy numerically and approximate the entropy for one- and two-parameter variants of the expanding map.

## Contents

# 1 Introduction

## 1.1 Defining the Circle Space

The ***circle***, a geometric figure whose locus is the set of all points in a plane equidistant from a center point, may be interpreted along the real number line as $\mathbb{S}^1 = [0, 1)$. In such a case, 0 acts as a position marker for some point on the circle and each $p \in \mathbb{S}^1$ represents a point on the circle whose clockwise distance from 0 is $2\pi r \cdot p$ where $r$ is the radius of the circle. Since it is possible to wrap around a circle infinitely many times, $p$ could actually represent all points along the circle whose clockwise distance from 0 is of the form $2\pi r \cdot (k + p)$ for all $k \in \mathbb{Z}$. Thus we can map all points in $\mathbb{R}$ onto $\mathbb{S}^1$ using the function $x \mapsto x \mod 1$. As the unit circle can also be represented using complex numbers, a similar definition of the circle in the complex plane is $\mathbb{S}^1 = \{e^{2\pi i z} \mid z \in [0, 1)\}$.

## 1.2 The Expanding Map on $\mathbb{S}^1$

The ***expanding map on*** $\mathbb{S}^1$ is a dynamical system with phase space $\mathbb{S}^1$ that represents an expansion of the number line along the circle. That is, if we were to consider:

1. Wrapping an elastic band around the circle,
2. Marking some point $p \in \mathbb{S}^1$ on the elastic band,
3. Stretching the elastic band to $k > 1$ times its length, and
4. Rewrapping the stretched elastic band around the circle

then the mapping from $p$ to its new location $p' \in \mathbb{S}^1$ would be the expanding map. We formally define this map as $f : \mathbb{S}^1 \to \mathbb{S}^1$ where $f(x) = kx \mod 1$ for some $k > 1$. In this thesis, we only focus on the case where $k = 2$ which signifies doubling the length of the elastic band, and any further reference to $f(x)$ specifically pertains to this case. A plot of $f(x)$ may be seen in Figure 1. It is

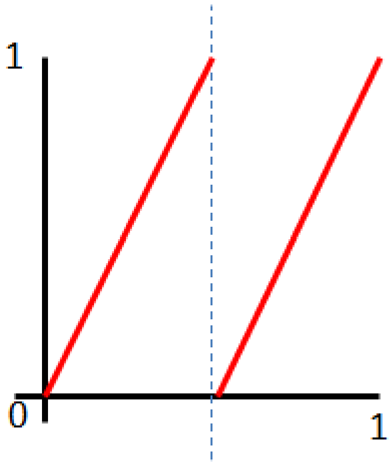


Figure 1: A plot of the expanding map on $\mathbb{S}^1$ $f(x)$

evident by this figure that $f(x)$ is invertible on small intervals whose cardinalities are at most half. This property classifies $f(x)$ as a ***local diffeomorphism***.

There are other simple, yet interesting traits of the expanding map which make it a popular choice for studying chaotic systems. To be a ***chaotic system***, the dynamical system must satisfy the following properties [1]:

- The map must be ***transitive***. This is satisfied if there exists an orbit generated by the map that is dense in the phase space.
- The set of all periodic points must be dense in the phase space.
- The map must have ***sensitive dependence on the initial conditions***. That is, two input values that are "nearby" are not guaranteed to yield output values that are "nearby."

The expanding map on $\mathbb{S}^1$ satisfies this definition by the following examples:

- The function can be shown to be transitive by the orbit generated by any irrational initial value. The orbit

  $$\pi \mod 1, 2\pi \mod 1, 3\pi \mod 1, 4\pi \mod 1, \ldots, k\pi \mod 1, \ldots$$

  has values that may be found between any two real numbers $x_1, x_2 \in \mathbb{S}^1$.
- The set of periodic points in $f$ is also dense in $\mathbb{S}^1$ since any $p \in \mathbb{S}^1$ whose least significant digit is even is periodic. Some simple examples are the orbits $0.2, 0.4, 0.8, 1.6 \mod 1 = 0.6, 1.2 \mod 1 = 0.2$ and $0.12, 0.24, 0.48, 0.96, 1.92 \mod 1 = 0.92, 1.84 \mod 1 = 0.84, 1.68 \mod 1 = 0.68, 1.36 \mod 1 = 0.36, 0.72, 1.44 \mod 1 = 0.44, 0.88, 1.76 \mod 1 = 0.76, 1.52 \mod 1 = 0.52, 1.04 \mod 1 = 0.04, 0.08, 0.16, 0.32, 0.64, 1.28 \mod 1 = 0.28, 0.56, 1.12 \mod 1 = 0.12$.
- We prove the sensitive dependence on initial conditions by comparing $0.5 + \epsilon$ and $0.5 - \epsilon$ which are only a distance of $2\epsilon > 0$ apart. As $\epsilon$ approaches 0, $0.5 + \epsilon$ approaches $2(0.5 + \epsilon) \mod 1 = 2\epsilon$ while $0.5 - \epsilon$ approaches $2(0.5 - \epsilon) \mod 1 = 1 - 2\epsilon$ so that the outputs are $1 - 4\epsilon$ apart. Hence decreasing the distance between the inputs through a smaller $\epsilon$ increases the distance between the respective outputs.

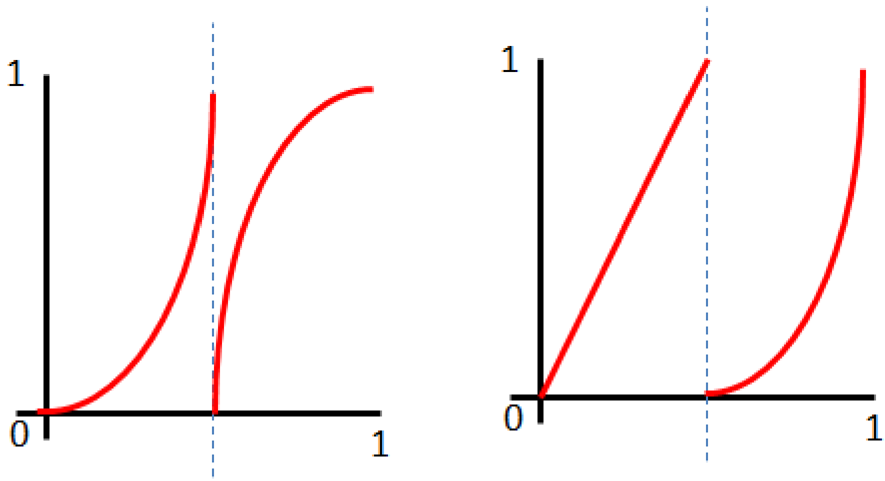


Figure 2: Plots of $g_*$ where $* = \epsilon < 0$ (Left) and $* = 0, \epsilon_2 > 0$ (Right)

### 1.3 Applying Perturbation to the Expanding Map $f$

As the expanding map $f(x)$ has been studied before, we choose to look at variations that preserve the chaotic system. De La Llave, Shub, and Simó have done similar work looking at variations of the form $x \mapsto kx + \alpha + \epsilon \sin(2\pi x)$. This exposed the expanding map to horizontal translations ($\alpha$), scalar dilations ($k$), and increased curvature ($\epsilon$) [4]. The increased curvature of the function results in a non-uniform expansion of the elastic band - some regions of the band will be stretched more than others. Our variations will focus on the changes in curvature in the expanding map using $k$-ary parameters. That is, for a list of $k$ parameters $* = \epsilon_1, \epsilon_2, \ldots, \epsilon_k$, we define the function $g_* : \mathbb{S}^1 \to \mathbb{S}^1$ as $g_*(x) = 2x + \epsilon_y \sin(2\pi x)$ where $y = z \in \{1, 2, \ldots, k\}$ if $\frac{z-1}{k} \leq x < \frac{z}{k}$. Our focus will be on the single-parameter variation $g_\epsilon(x) = 2x + \epsilon \sin(2\pi x)$ and the two-parameter variation $g_{\epsilon_1,\epsilon_2} = 2x + \epsilon_y \sin(2\pi x)$ where $y = 1$ if $0 \leq x < \frac{1}{2}$ and $y = 2$ if $\frac{1}{2} \leq x < 1$. Allowing more parameters would run the risk of breaking the continuity along the intervals $[0, 0.5)$ and $(0.5, 1)$. Examples of some of these variations may be viewed in Figure 2.

### 1.4 Invariant Measures on $g_*$

A ***measure*** is a function on a space that maps a value to its respective magnitude within the space. That is, it measures the value with regards to the definition of the space. For example, numbers in $p$-addic space with larger denominators have a greater measure than those with smaller denominators which gives $\frac{5}{32}$ a greater measure than 100. For this thesis, we will assume that all our measures have ***density functions***; $\rho$ is a density function along an interval

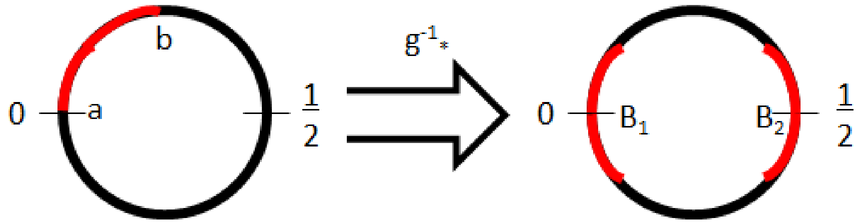


Figure 3: The inverse of the expanding map on $\mathbb{S}^1$ on some interval $[a, b]$

$[a, b]$ if:

- $\rho(x) \geq 0$ for all $x \in [a, b]$
- $\int_a^b \rho(x)dx = 1$

A common example of a density function is any probability distribution found in statistics. The measure $\mu$ that has density function $\rho$ is defined by:

$$\mu(c_1, c_2) = \int_{c_1}^{c_2} \rho(x)dx \text{ where } c_1, c_2 \in [a, b]$$

.

A measure $\mu$ is ***invariant*** under a function if the measure of the function's preimage on any small interval or set is equal to the measure of the interval or set itself. This would imply that the function's preimage of set $\mathcal{S}$ is equivalent to $\mathcal{S}$. Mathematically, we represent $\mu$ as an invariant measure under $g_*$ if $\mu(g_*^{-1}(\mathcal{S})) = \mu(\mathcal{S})$ where $g_*^{-1}(\mathcal{S}) = \{x \mid g_*(x) \in \mathcal{S}\}$. When $\mathcal{S}$ is an interval of the form $[a, b]$ such that $(b-a) < \frac{1}{2}$, it is clear that $g_*(a, b) = B_1 \cup B_2$ where $B_1 \cap B_2 = \emptyset$. One of these disjoint sets represents values about $0+\kappa$ while the other set represents values about $\frac{1}{2} + \kappa$; $\kappa$ is an offset dependent of the parameters of $g_*$.

In particular, when we consider the expanding map on $\mathbb{S}^1$ without perturbation, we can easily see that $\mu$ having density function $\rho(x) = 1$ is invariant under $f = g_0 = g_{0,0}$. The disjoint sets for the above interval $[a, b]$ have cardinality $|B_1| = |B_2| = \frac{1}{2}(b - a)$ centered about 0 and $\frac{1}{2}$. Figure 3 provides a visual example of this preimage property.

Furthermore, for our various perturbed expanding maps on $\mathbb{S}^1$, the invariant meausures defined by density functions $\rho$ are equivalent to the ***transfer function*** (also referred to as the ***transfer operator***)

$$L_*(\rho(x)) = \frac{\rho(g_{*;1}^{-1}(x))}{g_*'(g_{*;1}^{-1}(x))} + \frac{\rho(g_{*;2}^{-1}(x))}{g_*'(g_{*;2}^{-1}(x))} \text{ where } g_{*;1}^{-1}(x) \leq g_{*;2}^{-1}(x)$$

. Here $g_{*;1}^{-1}(x) \cup g_{*;2}^{-1}(x) = g_*^{-1}(\{x\})$ since every $p \in \mathbb{S}^1$ has two points in its preimage. This equivalence primarily comes from

**Proposition 1** $\mu$ having probability density $\rho$ is invariant under $g_*$ if and only if $\rho$ is a fixed point of the transfer operator $L_*$.

**Proof:**

($\Leftarrow$) First let $\mu$ be invariant. Then

$$\mu(g_*^{-1}[a,b]) = \mu([a,b]) = \int_a^b \rho(x)dx$$

where $[a,b]$ is a small interval on $\mathbb{S}^1$. Then because the inverse of the expanding map on an interval yields two non-overlapping intervals, $g_*^{-1}([a,b]) = \left[g_{*;1}^{-1}(a), g_{*;1}^{-1}(b)\right] \cup \left[g_{*;2}^{-1}(a), g_{*;2}^{-1}(b)\right]$. So we derive

$$\begin{aligned}
\mu(g_*^{-1}([a,b])) &= \mu(\left[g_{*;1}^{-1}(a), g_{*;1}^{-1}(b)\right] \cup \left[g_{*;2}^{-1}(a), g_{*;2}^{-1}(b)\right]) \\
&= \mu(\left[g_{*;1}^{-1}(a), g_{*;1}^{-1}(b)\right]) + \mu(\left[g_{*;2}^{-1}(a), g_{*;2}^{-1}(b)\right]) \\
&= \int_{g_{*;1}^{-1}(a)}^{g_{*;1}^{-1}(b)} \rho(x)dx + \int_{g_{*;2}^{-1}(a)}^{g_{*;2}^{-1}(b)} \rho(x)dx
\end{aligned}$$

Now we define $z = g_*(x)$ which means $x = g_*^{-1}(z)$ and $dz = g_*'(x) \cdot dx$ which may be rewritten as

$$dx = \frac{dz}{g_*'(x)} = \frac{dz}{g_*'(g_*^{-1}(z))}$$

Since $g_*(g_{*;j}^{-1}(m)) = m$ for all $(j,m) \in \{1,2\} \times \mathbb{S}^1$, we perform a u-substitution to get

$$\int_{g_{*;1}^{-1}(a)}^{g_{*;1}^{-1}(b)} \rho(x)dx + \int_{g_{*;2}^{-1}(a)}^{g_{*;2}^{-1}(b)} \rho(x)dx$$

$$= \int_a^b \rho(g_{*;1}^{-1}(z)) \cdot \frac{dz}{g_*'(g_{*;1}^{-1}(z))} + \int_a^b \rho(g_{*;2}^{-1}(z)) \cdot \frac{dz}{g_*'(g_{*;2}^{-1}(z))}$$

which further simplifies to

$$\int_a^b \left( \rho(g_{*;1}^{-1}(z)) \cdot \frac{1}{g_*'(g_{*;1}^{-1}(z))} + \rho(g_{*;2}^{-1}(z)) \cdot \frac{1}{g_*'(g_{*;2}^{-1}(z))} \right) dz$$

Recall that from this chain of equalities

$$\int_a^b \left( \frac{\rho(g_{*;1}^{-1}(z))}{g_*'(g_{*;1}^{-1}(z))} + \frac{\rho(g_{*;2}^{-1}(z))}{g_*'(g_{*;2}^{-1}(z))} \right) dz = \mu(g_*^{-1}([a,b])) = \mu([a,b]) = \int_a^b \rho(x)dx$$

As this applies for any arbitrarily selected $[a,b] \in \mathbb{S}^1$, we may conclude that

$$\rho(x) = \frac{\rho(g_{*;1}^{-1}(x))}{g_*'(g_{*;1}^{-1}(x))} + \frac{\rho(g_{*;2}^{-1}(x))}{g_*'(g_{*;2}^{-1}(x))} = L_*(\rho(x))$$

Thus $\rho$ is a fixed point of the transfer operator $L_*$.

($\Rightarrow$) Let $\rho$ be a fixed point of the transfer operator $L_*$. Then

$$\rho(x) = L_*(\rho(x)) = \frac{\rho(g_{*;1}^{-1}(x))}{g'_*(g_{*;1}^{-1}(x))} + \frac{\rho(g_{*;2}^{-1}(x))}{g'_*(g_{*;2}^{-1}(x))}$$

Integrating the left- and right-hand sides of the above equality from $a$ to $b$ with respect to $x$ gives us

$$\int_a^b \rho(x)dx = \int_a^b \frac{\rho(g_{*;1}^{-1}(x))}{g'_*(g_{*;1}^{-1}(x))} + \frac{\rho(g_{*;2}^{-1}(x))}{g'_*(g_{*;2}^{-1}(x))}dx$$

The left-hand side is clearly equal to $\mu([a,b])$. By performing the simplifications and u-substitutions from the $\Leftarrow$ component of this proof in reverse order, we can also simplify the right-hand side to $\mu(g_*^{-1}([a,b])$. Then we may conclude that $\mu([a,b]) = \mu(g_*^{-1}([a,b])$ by the transitive property of equality. As this lattermost equality satisfies the definition of invariance, we may conclude that $\mu$ is invariant under $g_*$. $\square$

As a consequence of Proposition 1, we will work with density functions and the transfer operator in order to find invariant measures on $g_*$. See Section 1.5.3..

## 1.5 Entropy of Dynamical Systems

For chaotic dynamical systems such as $f$ and $g_*$, it can be hard to understand their orbits. However, other dynamical systems such as the Leibniz Butterfly have orbits that are far more difficult to interpret. Thus ***entropy*** is a quantity that characterizes the system's complexity level.

In order to properly explain the computation of the entropy, we will first need to discuss measurable partitions. ***Measurable partitions*** are collections of measurable subsets that fully divide up an entire space and share no common elements. Each subset is considered to be ***measurable*** with respect to some function if it's inverse is also measurable. The use of the inverse in this definition acts recursively so that subset $\mathcal{C}_\alpha$ is measureable with respect to a function $h$ if the subsets $h^{-1}(\mathcal{C}_\alpha)$, $h^{-2}(\mathcal{C}_\alpha)$, $h^{-3}(\mathcal{C}_\alpha)$, ... are all measurable. For example, a measurable partition of $\mathbb{S}^1$ with respect to function $g_*$ is $\xi = \{\mathcal{C}_1, \mathcal{C}_2\} = \{[0, \frac{1}{2}), [\frac{1}{2}, 1)\}$.

To determine the entropy of a dynamical system, we compute the following for some measurable partition $\xi$ on measure $\mu$:

$$H(\xi) = H_\mu(\xi) = -\sum_{\alpha\in\xi} \mu(\mathcal{C}_\alpha) \log \mu(\mathcal{C}_\alpha)$$

where we allow $0 \cdot \log 0 = 0$. Those familiar with statistics and some fields of engineering may recognize this as Shannon's equation for entropy of information loss over a probability distribution,

$$H = -\sum_{i=1}^{n} p_i \log(p_i) \text{ where } \sum_{i=1}^{n} p_i = 1$$

and it is derived from the same concept. [2]

### 1.5.1 Entropy of Measure-Preserving Transformations

Given a measure-preserving transformation $T$, we extend the definition of a measurable partition $\xi = \{\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3, \ldots, \mathcal{U}_n\}$ to a joint partition. A ***joint partition*** $\xi^T_{-n}$ is a set that consists of all possible intersections of the partition subjected to the last $n-1$ iterations of the inverse of $T$. That is:

$$\xi^T_{-n} = \xi \vee T^{-1}(\xi) \vee T^{-2}(\xi) \vee \ldots \vee T^{-n+1}(\xi)$$

**Example:** Let us begin with a simple partition of $\mathbb{S}^1$ where we break it into two equal-sized intervals: $\xi = \{\mathcal{U}_1, \mathcal{U}_2\} = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$. We recall that $0 \equiv 1 \pmod 1$ when defining these intervals. Since $f$ is a measure-preserving transformation, we will compute the joint partition $\xi^f_{-2} = \xi \vee f^{-1}(\xi)$ where

$$\begin{aligned} f^{-1}(\mathcal{U}_1) = f^{-1}\left(\left[0, \tfrac{1}{2}\right]\right) = \left[0, \tfrac{1}{4}\right] \cup \left[\tfrac{3}{4}, 1\right] \\ f^{-1}(\mathcal{U}_2) = f^{-1}\left(\left[\tfrac{1}{2}, 1\right]\right) = \left[\tfrac{1}{4}, \tfrac{1}{2}\right] \cup \left[\tfrac{1}{2}, \tfrac{3}{4}\right] \end{aligned}$$

Since the binary operator $\vee$ yields the set of all intersections of elements in the left and right operands, $\xi^f_{-2}$ will contain four elements since $\xi$ and $f^{-1}(\xi)$ each contain two elements:

1. $\mathcal{U}_1 \cap f^{-1}(\mathcal{U}_1) = \left[0, \frac{1}{4}\right]$
2. $\mathcal{U}_1 \cap f^{-1}(\mathcal{U}_2) = \left[\frac{3}{4}, 1\right]$
3. $\mathcal{U}_2 \cap f^{-1}(\mathcal{U}_1) = \left[\frac{1}{4}, \frac{1}{2}\right]$
4. $\mathcal{U}_2 \cap f^{-1}(\mathcal{U}_2) = \left[\frac{1}{2}, \frac{3}{4}\right]$

The joint partition is used in computing the entropy with respect to $T$. Referred to as the ***metric entropy of*** $T$ ***relative to*** $\xi$, it is computed by $h(T, \xi) = h_\mu(T, \xi) = \lim_{n\to\infty} n^{-1} H(\xi^T_{-n})$ where $\mu$ is an invariant measure on $T$. $h(T, \xi)$ is guaranteed to exist, and it follows that the entropy of the measure-preserving transformation with respect to the measure is $h(T) = h_\mu(T) = \sup\{h_\mu(T, \xi) \mid \xi \text{ is a measurable partition of finite entropy}\}$. [2] The above supremum is reached when $\xi$ is a "generator partition" which is defined in [2].

**Example** Let us continue to use the partition $\xi = \left\{\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$ and measure-preserving transformation $f$ from the previous example. Katok and Hasselblatt state that $\xi$ is a generator partition for $f$ and measure $\mu$ where $\mu(\mathcal{S})$ is the length of the arc $\mathcal{S} \in \mathbb{S}^1$. Clearly $\xi = \xi^f_{-1}$, the cardinality of $\xi^f_{-1}$ is 2, and $\mu(\left[0, \frac{1}{2}\right]) = \mu(\left[\frac{1}{2}, 1\right]) = \frac{1}{2}$. By the results of the previous example, we also know that the cardinality of $\xi^f_{-2}$ is 4 and $\mu(\mathcal{U}) = \frac{1}{4}$ for all $\mathcal{U} \in \xi^f_{-2}$.

It can easily follow from induction that the cardinality of $\xi^f_{-n}$ is $2^n$ and $\mu(\mathcal{U}) = \frac{1}{2^n}$ for all $\mathcal{U} \in \xi^f_{-n}$. Thus we can compute:

$$\begin{aligned} H(\xi^f_{-n}) &= H_\mu(f, \xi) = -\textstyle\sum_{\mathcal{U} \in \xi^f_{-n}} \mu(\mathcal{U}) \log \mu(\mathcal{U}) = -2^n \cdot \left(\frac{1}{2^n} \cdot \log\left(\frac{1}{2^n}\right)\right) \\ &= -(-n)\log(2) = n\log(2) \end{aligned}$$

It follows that $h_\mu(f,\xi) = \lim_{n\to\infty} n^{-1} H_\mu(f,\xi) = \lim_{n\to\infty} n^{-1}\cdot(n\log(2)) = \log(2)$ is the metric entropy of $f$ relative to $\xi$.

**Proposition 2** For the measure-preserving expanding maps on $\mathbb{S}^1$ of the form $j(x) = kx \pmod 1$, $h_\mu(j) = \log(k)$ for all $k \in \mathbb{N}$.

### 1.5.2 Topological Entropy

Proposition 2 now allows us to find the entropy $h_\mu(f)$ as a function with respect to the measure $\mu$ alone. However, we still need to have some caution since there exist many invariant measures for a given transformation $T$. For example, the point-mass measure at $\{0\}$ defined for $\mathcal{S} \subset \mathbb{S}^1$ as:

$$\mu(\mathcal{S}) = \begin{cases} 1 \text{ if } 0 \in \mathcal{S} \\ 0 \text{ if } 0 \notin \mathcal{S} \end{cases}$$

is invariant with respect to the expanding map $f$ since $\mu(f^{-1}(\mathcal{S})) = \mu(\mathcal{S})$ for all possible $\mathcal{S}$. This follows from the fact that 0 is a fixed point of $f$ and that $0 \equiv 1 \pmod 1$. As this maps all subsets of $\mathbb{S}^1$ to the fixed point, we conclude that the entropy is $h_\mu(f) = \log(1) = 0$ in this case.

Thus the entropy of a measure-preserving transformation $T$ varies with respect to the chosen measure $\mu$. This arises from the fact that the entropy with some measure $\mu$ is the observed complexity through that particular measure. Since a measure is nothing more than a tool to provide perspective relations to the elements in a space, the observation of the transformation's entropy is limited to the individual measure's viewpoint of the space.

If we prefer to remove this bias of perspective, then we must consider a view where measures see all elements of the space equally (that is, where $\mu(x) = C$ for all $x$ in the space). The ***topological entropy*** is defined topologically to accomplish this so that no measure is involved in the observation. The topological entropy may be determined by the theorem below.

**Theorem 1** For a fixed measure-preserving transformation $T$, the topological entropy is $h_\mu(T) = \sup_\mu\{h_\mu(T)\}$.

**Example** For any expanding map $f$ on the circle space, the topological entropy happens to be $\log(n)$ where $n$ is the number $k$ that is the scalar of $x$ which we found in Proposition 2.

[2]

### 1.5.3 Weighted Entropy

Let us define a ***potential function*** $\varphi : \mathbb{S}^1 \to \mathbb{R}$ that is continuous. In particular, the purpose of $\varphi$ is to apply varying weights to the points along the circle space. Then the ***weighted entropy*** of a measure-preserving transformation $T$ with respect to $\varphi$ and invariant measure $\mu$ is

$$h_\mu(T) + \int \varphi d\mu$$

The invariant measure $\mu^*$ that yields the maximum value of the above formula for a fixed $T$ and $\varphi$ is referred to as the ***equilibrium measure/state of*** $\varphi$. We present some propositions below regarding the existence and uniqueness of $\mu^*$ and direct the reader to [5] for further details and proofs.

**Proposition 3** $\mu^*$ exists if $\varphi$ is continuous.

**Proposition 4** $\mu^*$ is unique if $\varphi$ is ***Hölder continuous***. This means that for all $\alpha > 0$, there exists a constant $C \in \mathbb{R}$ such that $|\varphi(x) - \varphi(y)| \leq C\,|x - y|^{\alpha}$.

We consider a special potential function with regards to $T$. It is $\varphi(x) = -\log(T'(x))$ which is continuous. Any example where $T'$ is Hölder continuous implies that $\varphi$ is Hölder continuous. Then we have satisfied Propositions 3 and 4 so that $\mu^*$ exists and is unique. For this particular potential function, $\mu^*$ is also called the ***Sinai-Ruelle-Bowen (SRB) measure*** or the ***physical measure*** since it captures the behavior of typical orbits of the system from the equilibrium's point-of-view. Due to capturing this behavior, it is the most important of the entropy observations.

We recall from Section 1.4 that as an invariant measure, $\mu^*$ has a density function $\rho$ on $\mathbb{S}^1$. That is, there exists a density function $\rho(x) \geq 0$ such that

$$\mu^*([a,b]) = \int_a^b \rho(x)dx$$

for all intervals $[a,b] \in \mathbb{S}^1$. $\mu^*$ is invariant with respect to $T$ if and only if the associated density function $\rho$ is the fixed point of the transfer operator $L_*$ for the measure-preserving transformation $T$.

Hence iteratively computing the measure on $T$ is equivalent to iteratively computing the transfer operator $L_*$ on the measure's density function. We can find $\mu^*$ by starting with any initial density function which we will call $\rho_0$. We note that $\rho_0$ does not need to be invariant. A new density function $\rho_1$ can then be induced by $L_*$ as $\rho_1(x) = L_*(\rho_0(x))$. Furthermore, we may induce later density functions by $L_*$ as $\rho_n(x) = L_*(\rho_{n-1}(x))$. As a fixed point (which happens to be an attractor), it follows that $\lim_{n\to\infty} L_*^n(\rho_0(x)) = \rho(x)$ associated with the SRB measure $\mu^*$. Upon finding the equilibrium state $\mu^*$, we can compute the entropy with respect to it by taking advantage of Theorem 2 below:

**Theorem 2** (Ruelle) The weighted entropy of $T$ with respect to $\mu^*$ when $\mu^*$ is the SRB equilibrium state is 0. [5]

Thus we can solve for $h_{\mu^*}(T)$ via the following integral computation:

$$\begin{gathered} h_{\mu^*}(T) + \textstyle\int \varphi \cdot d\mu^* = 0 \\ \Downarrow \\ h_{\mu^*}(T) + \textstyle\int (-\log(T'(x))) \cdot \rho(x)dx \\ \Downarrow \\ h_{\mu^*}(T) = \textstyle\int (\log(T'(x))) \cdot \rho(x)dx \end{gathered}$$

where $\rho(x)$ can be found through infinite iteration as shown above and $T'(x)$ can be computed using differential calculus. Due to the difficulty of performing an infinite number of iterations, we will use numerical methods to obtain a reasonable convergence estimate of $\rho(x)$. We will also use numerical methods to approximate the integral needed to find the ***SRB entropy*** $h_{\mu^*}(T)$ since it may be difficult to derive for particular choices of $T$.

## 1.6 Validity of Methodology

In Section 1.5.3., we provide a method for finding the SRB measure that involves infinitely iterating the transfer operator $L_*$ on an arbitrary initial density function $\rho_0$. However, the contraction mapping theorem from functional analysis does not apply since $L_*$ is not a contracting map. That is, we have no guarantee that

$$d(\rho_i(a), \rho_i(b)) > d(\rho_{i+1}(a), \rho_{i+1}(b))$$

which means that each density function is not guaranteed to shrink towards the fixed point.

Depsite this, it can easily be shown that $L_*$ is a ***linear operator*** such that $L_*(\rho(x) + \eta(x)) = L_*(\rho(x)) + L_*(\eta(x))$ and $L_*(c \cdot \rho(x)) = c \cdot L_*(\rho(x))$ for all density functions $\rho(x)$ and $\eta(x)$ as well as all scalars $c \in \mathbb{R}$. Thus we are able to study the spectrum of $L_*$ and it happens to be the case that its largest eigenvalue is $\lambda = 1 \Rightarrow L_*(\rho(x)) = \rho(x)$. For $L_*$ to have been a contracting map, it would have been necessary for the absolute value of all its eigenvalues to be strictly less than 1.

Yet there exists a ***gap*** between this eigenvalue of 1 and the remaining eigenvalues of $L_*$. This gap is defined by the existence of a real number $\lambda_0 < 1$ such that $\sup\{|\lambda| \,|\lambda \in \Lambda(L_*) \setminus 1\} = \lambda_0$ where $\Lambda(L_*)$ is the collection of all eigenvalues of $L_*$. An illustration of this gap may be viewed in Figure 4. We further note that the ***Lasota-Yorke inequalities*** are satisfied; we refer the reader to [3] for an explanation of this condition.

Lasota has also shown that when a map $P(x)$ is not contracting, we can still guarantee convergence to the fixed point of the map $x^*$ through infinite iteration as long as:

1. The greatest eigenvalue of $P$ is 1.
2. There is a gap between the eigenvalue of 1 mentioned above and all other eigenvalues.
3. The Lasota-Yorke inequalities are satisfied.

That is, given that the three conditions are satisfied,

$$\lim_{n \to \infty} P^n(x) = x^*$$

Because $L_*$ satisfies the conditions, it also follows that $L_*(\rho_0) \to \rho(x)$ where $\mu^*([a, b]) = \int_a^b \rho(x)dx$. [3]

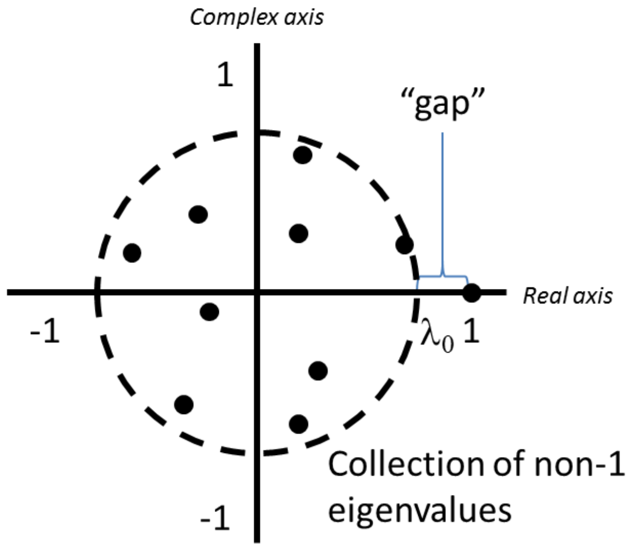


Figure 4: A graphical representation of a gap between the eigenvalue of 1 and the remaining eigenvalues in the complex plane. The supremum is represented by the dashed circle with radius $\lambda_0$.

# 2 Methods Used in Numerical Computation

## 2.1 Numerically Computing the Density function $\rho(x)$

We recall that $L_*$ is the transfer operator defined with respect to $g_*$ as:

$$L_*(\rho(x)) = \frac{\rho(g^{-1}_{*;1}(x))}{g'_*(g^{-1}_{*;1}(x))} + \frac{\rho(g^{-1}_{*;2}(x))}{g'_*(g^{-1}_{*;2}(x))}$$

where $g_*(x) = (2x + \epsilon_i \sin(2\pi x)) \mod 1$ if $\frac{i-1}{k} \leq x < \frac{i}{k}$. $k$ is the number of parameters provided in the form of $* = \epsilon_1, \epsilon_2, \ldots, \epsilon_k$. As $g_*$ is a fixed function, we can precompute the inverses and first derivatives needed for computing $L_*$. Then we only have to find the value of the density function at the inverse values as we compute each iteration of $L_*$; the remaining addition and division computations are trivial operations.

### 2.1.1 Precomputation Using $g_*$

Upon selecting our parameters $*$, we can easily find the first derivative of $g_*$ using elementary differential calculus:

$$g'_*(x) = 2 + 2\pi\epsilon_i \cos(2\pi x) \text{ if } \frac{i-1}{k} \leq x < \frac{i}{k}$$

Likewise, we can approximate the two inverse values of $g_*^{-1}$ using the Newton-Raphson Method. That is, we will minimize $||x - g_*(\hat{y})||_1$ by iteratively computing

$$\hat{y}_{n+1} = \hat{y}_n - \frac{g_*(\hat{y}_n)}{g'_*(\hat{y}_n)}$$

with initial values $\hat{y}_0 = 0.25$ to determine $g_{*;1}^{-1}(x)$ and $\hat{y}_0 = 0.75$ to determine $g_{*;2}^{-1}(x)$. We set our threshold for the numerically converged stopping condition at $10^{-9}$. Due to the limitations of computer memory, we cannot obtain these precomputations for every $x \in [0, 1]$ (memory is a countably finite resource). Thus we will perform these computations at every $x \in \{0, \frac{1}{j}, \frac{2}{j}, \dots, \frac{j-1}{j}, 1\}$ for some $j \in \mathbb{N}$. While greater $j$ require more memory and time for precomputation, it makes the available interval more fine-grained which improves our accuracy.

### 2.1.2 Numerical Implementation Details for Finding $\rho(x)$

Due to the discrete nature of computers, we are limited to numerical methods that only approximate each $\rho_i(x)$. In particular, we are only able to compute $\rho_i(x)$ for a specific $x$ if $x$ is one of the selected points along the interval $[0, 1]$ used in computation. All input values between two consecutive selected points must be approximated or "ignored." This leads to questioning if $\{\rho_i(x)\}$ ever converges to $\phi_\epsilon(x)$ as it would in continuous space. If it does, we could follow this with the question of whether or not the convergence would be quicker since only a subset of points along the interval $[0, 1]$ have to converge. Clearly such an issue only applies to point-wise convergence since uniform convergence applies to every point in the interval.

There is also the issue of determining which discrete representation we will choose to use. Each one will yield a different value of $\rho_i(x)$ for a given $x$. One method uses just the selected points and rounds $x$ to the nearest selected point (***endpoint method***) and the other is to proportionally average the result with respect to the two selected points between which $x$ is found (***interpolation method***). That is, the interpolation method approximates along the line segment $\rho_i(x) \approx \frac{x_2 - x}{x_2 - x_1}\rho_{i-1}(x_1) + \frac{x - x_1}{x_2 - x_1}\rho_{i-1}(x_2)$ where $x_1 \leq x \leq x_2$.

Regardless of our approximation method, we will often overestimate or underestimate the values of $\rho_i(x)$ which will prevent the density function from satisfying its property of representing a distribution along $[0, 1]$. That is, it may not be the case that

$$\int_0^1 \rho_i(x)dx = 1$$

after using the endpoint or interpolation methods. However, the proportionality amongst values still exists so that we may normalize the density function. Our normalization constant will simply be the area under the density function along the interval $[0, 1]$. Because we only have access to the discretized interval, we will need to use a Riemann sum to approximate this interval as well. Due to its greater accuracy, though, we will use Simpson's Rule for approximation instead.

Thus we numerically compute the normalization constant as

$$Z = \frac{1}{3n}\left[\rho_i(0) + 2\sum_{m=1}^{\frac{n}{2}-1}\rho_i\left(\frac{2m}{n}\right) + 4\sum_{m=0}^{\frac{n}{2}-1}\rho_i\left(\frac{2m+1}{n}\right) + \rho_i(1)\right]$$

and then normalize $\frac{\rho_i(x)}{Z} \to \rho_i(x)$ where $\to$ is the assignment operator. [6]

## 2.2 Numerically Computing the SRB Entropy $h_{\mu^*}(g_*)$

Upon approximating the density function $\rho(x)$ that is associated with the SRB measure, we can numerically calculate the SRB entropy

$$h_{\mu^*}(g_*) = \int_0^1 (\log(g'_*(x))) \cdot \rho(x)dx$$

We still have the precomputed derivative described in Section 2.1.1, and most modern programming languages contain an implementation of the logarithm function in their mathematics package. The remaining multiplication between the log of the derivative and the density is a trivial operation. To get the best approximation of this integral, we again use Simpson's Rule

$$\frac{1}{3n}\left[\log(g'_*(0)) \cdot \rho(0) + 2\sum_{m=1}^{\frac{n}{2}-1}\log\left(g'_*\left(\frac{2m}{n}\right)\right) \cdot \rho\left(\frac{2m}{n}\right) + \right.$$

$$\left. 4\sum_{m=0}^{\frac{n}{2}-1}\log\left(g'_*\left(\frac{2m+1}{n}\right)\right) \cdot \rho\left(\frac{2m+1}{n}\right) + \log(g'_*(1)) \cdot \rho(1)\right]$$

[6] Like any discrete approximation over a continuous interval, we obtain a more accurate estimate as our number of discrete components increases to make the interval more fine-grained.

# 3 Discussion

## 3.1 Results and Conjectures

After implementing the pseudocode from Section 2.3 in the C/C++ programming language, we were able to run the program with various choices for the parameter list $*$. However, because each parameter is restricted to the interval $\epsilon_i \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$, we are able to plot the SRB entropy over the entire set of parameter lists in one and two dimensions. These plots may be viewed in Figures 5 and 6, respectively. We recall that the one-dimensional version of $g_*$ is equal to the two-dimensional versions where $\epsilon_1 = \epsilon_2$. So Figure 5 can also be considered a cross-section of Figure 6 taken through the $\epsilon_1 = \epsilon_2$ plane.

As expected from our example in Section 1.5.1, the entropy at $\epsilon_1 = \epsilon_2 = 0$ is $\log(2)$ which is the entropy of $f$. We may also observe from the contour plots in

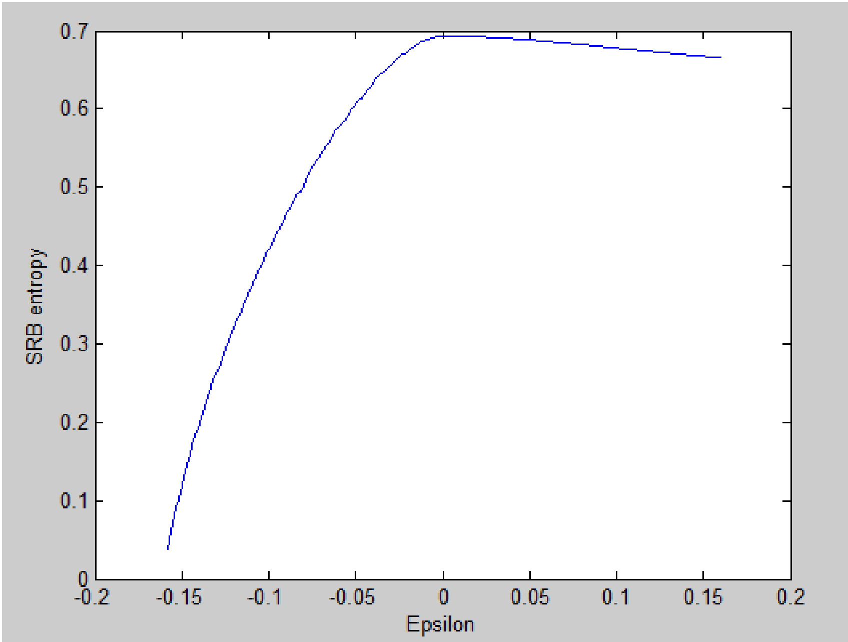


Figure 5: Plot of the SRB Entropy for $g_\epsilon(x)$ along the interval $\epsilon \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$.

Figure 7 that the SRB entropy seems to strictly decrease as the parameters approach the endpoints of the intervals. Thus we present the following conjectures regarding the entropy of these variations on the expanding map:

**Conjecture 1** For all curves defined from $\epsilon_1 = \epsilon_2 = 0$ to either $\epsilon_1 = -\frac{1}{2\pi}, \epsilon_2 \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$ or $\epsilon_1 \in [-\frac{1}{2\pi}, \frac{1}{2\pi}], \epsilon_2 = \pm\frac{1}{2\pi}$, $h_\mu(g_{\epsilon 1, \epsilon 2})$ is monotonically decreasing.

**Conjecture 2**

$$\inf \left\{ h_\mu \left( g_{\epsilon_1, \epsilon_2} \right) \middle| \epsilon_1, \epsilon_2 \in \left[ -\frac{1}{2\pi}, \frac{1}{2\pi} \right] \right\} = 0$$

We also found the equilibrium density functions $\rho$ to have an interesting form. It appears that as the values of the parameters decrease, $\rho$ develops a more extreme exponential shape where the values near either 0 or 1 are large and the rest of the interval quickly decreases towards extrememly small values. On the other hand, an increase in the values of the parameters leads to a more polynomial or trigonometric-shaped curve for $\rho$. The curves are mostly asymmetrical in form so that it is neither the case that $\rho(x) = \rho(1-x)$ nor that $\rho'(x) = -\rho'(1-x)$. Examples of these density functions may be seen in Figure 8.

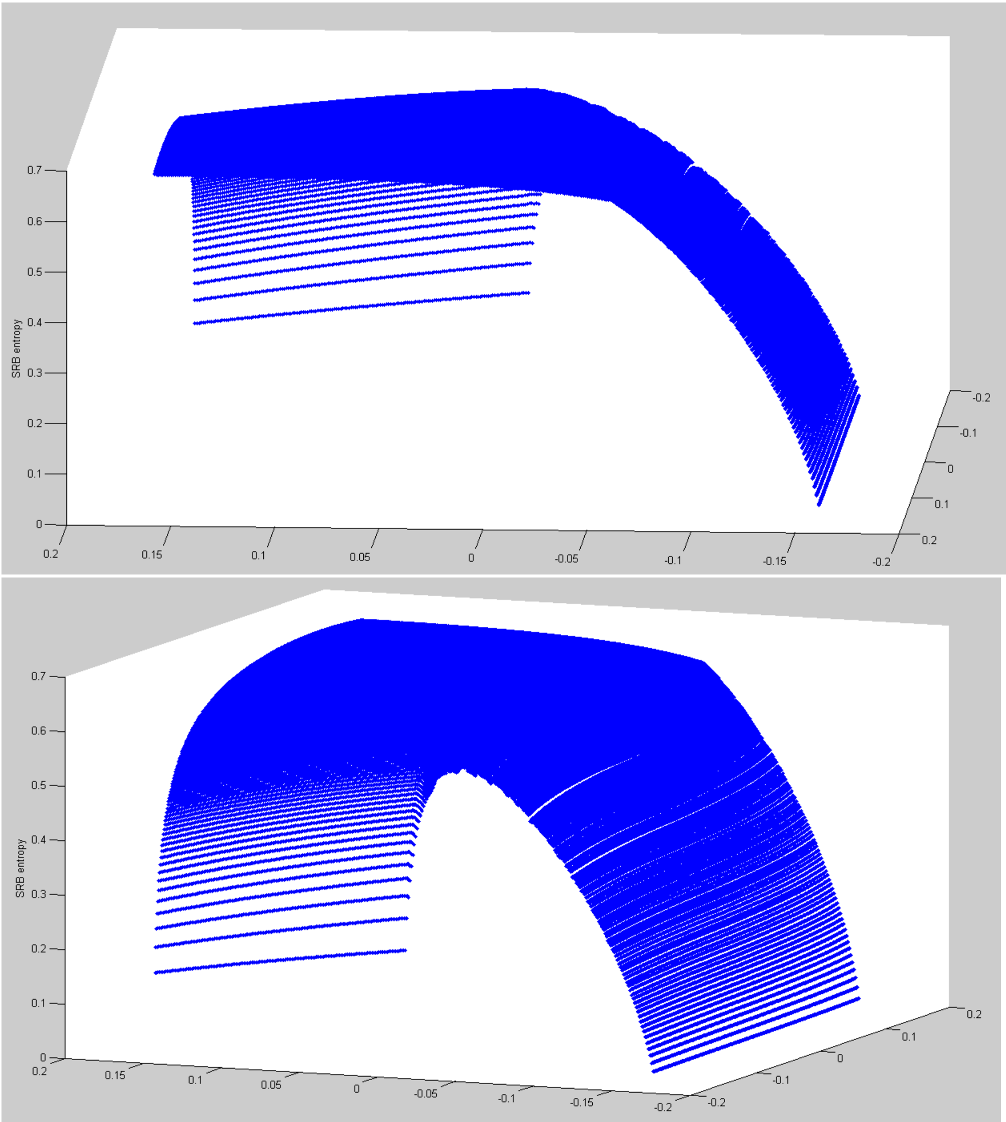


Figure 6: Plot of the SRB Entropy for $g_{\epsilon_1,\epsilon_2}(x)$ along the interval $\epsilon_1, \epsilon_2 \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$ from two perspectives.

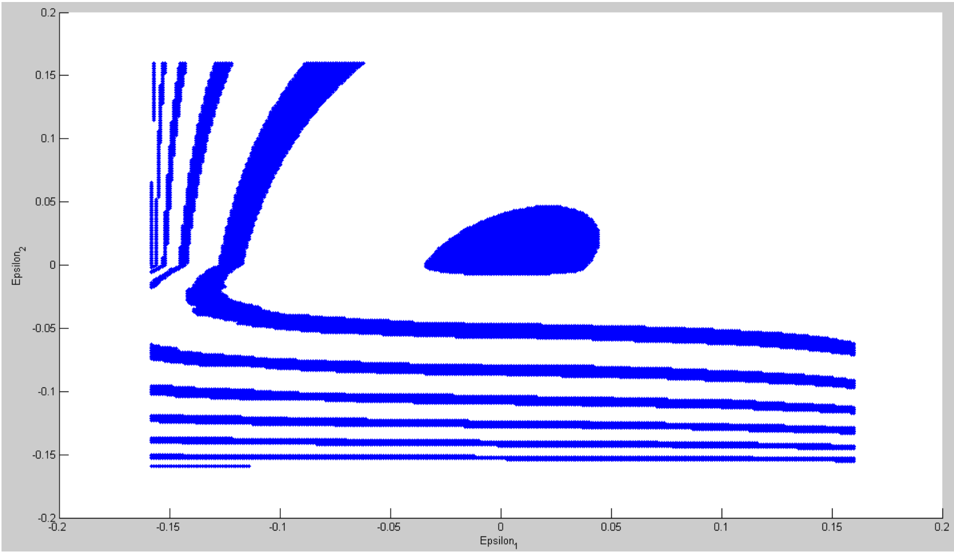


Figure 7: Contour plot of the SRB Entropy for $g_{\epsilon_1,\epsilon_2}(x)$ along the interval $\epsilon_1, \epsilon_2 \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$.

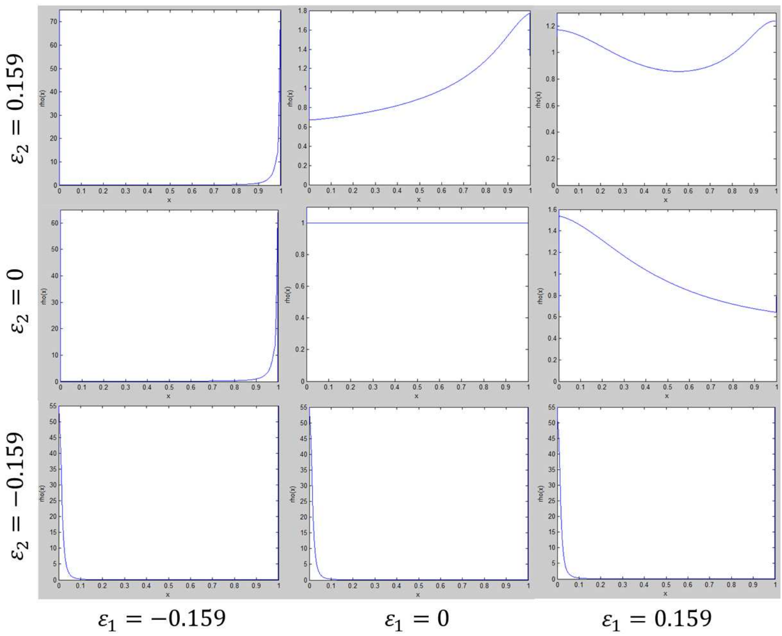


Figure 8: Examples of the density function $\rho$ with respect to specific $g_{\epsilon_1,\epsilon_2}(x)$ along the interval $\epsilon_1, \epsilon_2 \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$.

### 3.2 Future Research

As we only approximated the entropies of the perturbations of the expanding map $g_*$ through numerical computations, we plan to use analytic techniques and theory in order to study the validity of our conjectures above. Furthermore, we hope to study the variation of the SRB Entropy of $g_*$ with more parameters. Due to the risk of losing continuity with these additional parameters when applying perturbation to $f(x) = 2x \mod 1$, we would require that $*$ has at most $k$ parameters $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$ when $g_*$ corresponds to the perturbation of $f(x) = kx \mod 1$. That is, we will consider more parameters when we increase the rate of expansion.

## References


[1] Elaydi, Saber N. *"Discrete Chaos Second Edition: With Applications in Science and Engineering",* Taylor & Francis Group, 2008.

[2] Katok, Anatole and Hasselblat, Boris. *"Introduction to the Modern Theory of Dynamical Systems",* Cambridge University Press, 1995.

[3] Lasota A. and Mackey, M. C. *"Chaos, Fractals, and Noise Second, Edition",* Springer-Verlag, New York, 1995.

[4] Llave, Rafael De La, Shub, Michael and Simó, Carles. *Entropy Estimates for a Family of Expanding Maps of the Circle,* Discrete and Continuous Dynamical Systems, Vol. 10, pp 597-608, 2008.

[5] Ruelle, David. *"Thermodynamic Formalism, Encyclopedia of Mathematics and Its Applications, Volume 5",* Addison-Wesley, London, 1978.

[6] Weisstein, Eric W. "Simpson's Rule." From *MathWorld* – A Wolfram Resource. *http://mathworld.wolfram.com/SimpsonsRule.html*


# 4 Appendix A: List of Functions and Variables Used

$f(x) = kx \mod 1$ for any $k \in \mathbb{N} \setminus 1$ (*Expanding Map*)

$*$ is a list of $k \in \mathbb{N}$ parameters such that $\epsilon_1, \epsilon_2, \ldots, \epsilon_k \in [-\frac{1}{2\pi}, \frac{1}{2\pi}]$

$g_*(x) = (kx + \epsilon_i \sin(2\pi x)) \mod 1$ if $\frac{i-1}{k} \leq x < \frac{i}{k}$ for any $k \in \mathbb{N} \setminus 1$

$\rho_n(x)$ is a density function on $[0, 1]$

$L_*(\rho(x)) = \frac{\rho(g_{*,1}^{-1}(x))}{g_*'(g_{*,1}^{-1}(x))} + \frac{\rho(g_{*,2}^{-1}(x))}{g_*'(g_{*,2}^{-1}(x))}$ where $g_{*,1}^{-1}(x) \leq g_{*,2}^{-1}(x)$ (*Transfer Function*)

$\rho(x)$ is the equilibrium density function on $[0,1]$ such that $L_*(\rho) = \rho$

$\mu$ is a measure

$T$ is a measure-preserving transformation in a given space

$\xi$ is a measurable partition of a space

$H_\mu(\xi) = -\sum_{\alpha\in\xi} \mu(\mathcal{C}_\alpha) \log \mu(\mathcal{C}_\alpha)$ where $0 \cdot \log 0 = 0$ (*Entropy of a dynamical system*)

$H = -\sum_{i=1}^{n} p_i \log(p_i)$ where $\sum_{i=1}^{n} p_i = 1$ (*Shannon Entropy for information loss*)

$\varphi$ is a potential function that applies weight along an interval

$h_\mu(T) + \int \varphi d\mu$ is the weighted entropy

When $\varphi(x) = -\log(T'(x))$, the invariant measure $\mu^*$ that yields the maximum weighted entropy is the *SRB measure*

$h_\mu(T) = \int_0^1 \log(T'(x)) \cdot \rho(x)dx$ (*Pesin's Theorem*)

# 5 Appendix B: Pseudocode

**Algorithm 1** Compute SRB Entropy $h_{\mu^*}(g_*)$

**return** $\int_0^1 (\log(g_*'(x)) \cdot \rho(x)) \cdot dx$; ▷ Simpson's Rule

**Algorithm 2** Approximate $g_*^{-1}(x)$ Using the Newton-Raphson Method

**for** $i$ **from** $0$ **to** $k-1$ **by** $1$ **do** ▷ $k$ = number of parameters in $*$
    $point \leftarrow \frac{x+i}{k}$; ▷ Initial point to test
    $difference \leftarrow \text{abs}(g_*(point) - x)$;
    **while** $difference > CONVERGENCETHRESHOLD$ **do**
        $point \leftarrow point - \frac{g_*(point)-x}{g_*'(point)}$;
        $difference \leftarrow \text{abs}(g_*(point) - x)$;
    **end while**
    $inverses[i] \leftarrow point$; ▷ Inverse point is found
**end for**
**return** $inverses$;

**Algorithm 3** Compute Equilibrium Density Function $\rho$

$\triangleright$ Precompute the values of $g_*^{-1}(x)$ and $g'_*(g_*^{-1}(x))$ for the transfer function
**for** $i$ **from** $0$ **to** $2 \cdot TOTALPOINTS$ **by** $1$ **do**
  $precompGInv[i] \leftarrow g_*^{-1}(\frac{i}{2 \cdot TOTALPOINTS})$; $\triangleright$ Newton-Raphson Method
  $precompGPrm[i][1] \leftarrow g'_*(precompGInv[i][1])$;
  $precompGPrm[i][2] \leftarrow g'_*(precompGInv[i][2])$;
**end for**
**while not** converged($\rho_{previous}$, $\rho_{current}$) **do**
  **for** $i$ **from** $0$ **to** $2 \cdot TOTALPOINTS$ **by** $1$ **do**
    $\rho_{previous}[i] \leftarrow \rho_{current}[i]$;
    $\rho_{current}[i] \leftarrow \frac{\rho_{current}[\lfloor precompGInv[i][1] \cdot 2 \cdot TOTALPOINTS \rfloor]}{precompGPrm[i][1]} + \frac{\rho_{current}[\lfloor precompGInv[i][2] \cdot 2 \cdot TOTALPOINTS \rfloor]}{precompGPrm[i][2]}$; $\triangleright$ Transfer Function
  **end for**
  $normalizeConstant \leftarrow \int_0^1 \rho_{current}(x) \cdot dx$; $\triangleright$ Simpson's Rule
  **for** $i$ from $0$ to $2 \cdot TOTALPOINTS$ by $1$ **do**
    $\rho_{current}[i] \leftarrow \frac{\rho_{current}[i]}{normalizeConstant}$;
  **end for**
**end while**
**return** $\rho_{current}$; $\triangleright$ Equilibrium density function $\rho$ obtained

**Algorithm 4** Approximate $\int_0^1 f(x) \cdot dx$ Using Simpson's Rule

$integratedValue \leftarrow f(0)$;
$\Delta \leftarrow \frac{1}{2 \cdot TOTALPOINTS}$;
**for** $i$ **from** $1$ **to** $(2 \cdot TOTALPOINTS) - 1$ **by** $2$ **do** $\triangleright$ $i$ is odd
  $integratedValue \leftarrow integratedValue + 4f(i \cdot \Delta) + 2f((i+1) \cdot \Delta)$;
**end for**
$\triangleright$ Only want one copy of $f(1)$ in the summation
$integratedValue \leftarrow integratedValue - f(1)$;
$integratedValue \leftarrow integratedValue \cdot \Delta \cdot \frac{1}{3}$;
**return** $integratedValue$;