Mostly Conjugate:
Relating Dynamical Systems — Beyond Homeomorphism

Joseph D. Skufca[1,*] and Erik M. Bollt[1,†]

[1] *Department of Mathematics, Clarkson University, Potsdam, New York*

(Dated: April 24, 2006)

A centerpiece of Dynamical Systems is comparison by an equivalence relationship called topological conjugacy. Current state of the field is that, generally, there is no easy way to determine if two systems are conjugate or to explicitly find the conjugacy between systems that are known to be equivalent. We present a new and highly generalizable method to produce conjugacy functions based on a functional fixed point iteration scheme. In specific cases, we prove that although the conjugacy function is strictly increasing, it is a.e. differentiable, with derivative 0 everywhere that it exists — a Lebesgue singular function. When applied to non-conjugate dynamical systems, we show that the fixed point iteration scheme still has a limit point, which is a function we now call a "commuter" — a non-homeomorphic change of coordinates translating between dissimilar systems. This translation is natural to the concepts of dynamical systems in that it matches the systems within the language of their orbit structures. We introduce methods to compare nonequivalent systems by quantifying how much the commuter functions fails to be a homeomorphism, an approach that gives more respect to the dynamics than the traditional comparisons based on normed linear spaces, such as L^2. Our discussion includes fundamental issues as a principled understanding of the degree to which a "toy model" might be representative of a more complicated system, an important concept to clarify since it is often used loosely throughout science.

PACS numbers: 05.45

Contents

*Electronic address: jskufca@clarkson.edu
†Electronic address: ebollt@clarkson.edu

1. INTRODUCTION

Since the beginnings of the field of dynamical systems by Henri Poincare' [1], the question characterizing a dynamical system has been to examine topological and geometric features of orbits, rather than focusing on the specific empirical details of the solutions of the dynamical system with respect to the details of the specific coordinate system. One seeks to understand coordinate independent properties of a dynamical system, such as the periodic orbit structure — the count and stability of periodic orbits. Specifically, the question of comparing two dynamical systems as dynamically the same has evolved into the modern notion to identify if there is a conjugacy between them [6, 8, 9, 11, 13].

While characterizing a dynamical system explicitly up to conjugacy is the central problem in areas ranging widely from bifurcation theory to studies of chaotic dynamical systems, to name a few, there are surprisingly few chaotic system pairs for which a conjugacy is known [2]. The primary, and usually the sole complete example, that one finds in textbooks is that between the logistic map, $x' = g_r(x) = rx(1-x)$, and the tent map, $y' = T_a(y) = a(1-2|x-\frac{1}{2}|)$, each as mappings of the unit interval. Only when each is two-to-one, that is $r = 4$, and $a = 2$, is the form of the conjugacy known: $h : [0,1] \to [0,1]$, where, $h(x) = \frac{1}{2}[1-\cos(\pi x)]$. While it may be known through other considerations, such as through the kneading theory, or comparing entropies, that two different versions of these mappings, g_r and T_a, with special values of r and a may be conjugate [5], the form of the conjugacy is generally not known. This is arguably an uncomfortable situation given that conjugacy is a centerpiece of dynamical systems theory and especially considering that these two are considered to be some of the simplest maps and are commonly used for benchmark studies. Our techniques allow us to generate conjugacies, such as the change of coordinates between two tent maps, illustrated in Fig. 1, shown in the form of a quadweb diagram, which we present in Sec. 11.

The common scenerio in dynamical systems is that we do not know if two given dynamical systems are equivalent (conjugate), but we may suspect conjugacy through investigations of the periodic orbit structures. Often, one may try to study the symbolic dynamics when a generating partition can be found (though finding such partitions is often a highly nontrivial task in higher-dimensional systems [3, 4]). Through comparison of one particularly popular figure of merit, the topological entropies, we may suspect that the two given systems are equivalent. However, even if the generating partitions are available, the equivalence between the maps and their symbolic dynamics is only a semiconjugacy [13] (one-one-ness is missing), and so there can be a loss of information when considering only the symbolic dynamics. This situation is commonly understood when studying a logistic map in its period-3 window, say $r = 3.84$. (Fig. 14.) There is a kneading sequence for this map which defines the grammar of its symbolic dynamics, and there is a value of a for the tent map with the same kneading sequence. However, that tent map is chaotic while the logistic map is chaotic only for a measure zero set of initial conditions on the chaotic saddle, with a.e. initial condition in the basin of the period-3 orbit. Symbolic dynamics is not definitive, nor is a measure derived from them, but they are suggestive. For these reasons, we consider there to be a considerable gap in the field of dynamical systems that for very few systems is there a direct and explicit comparison by conjugacy. The previous difficulty of

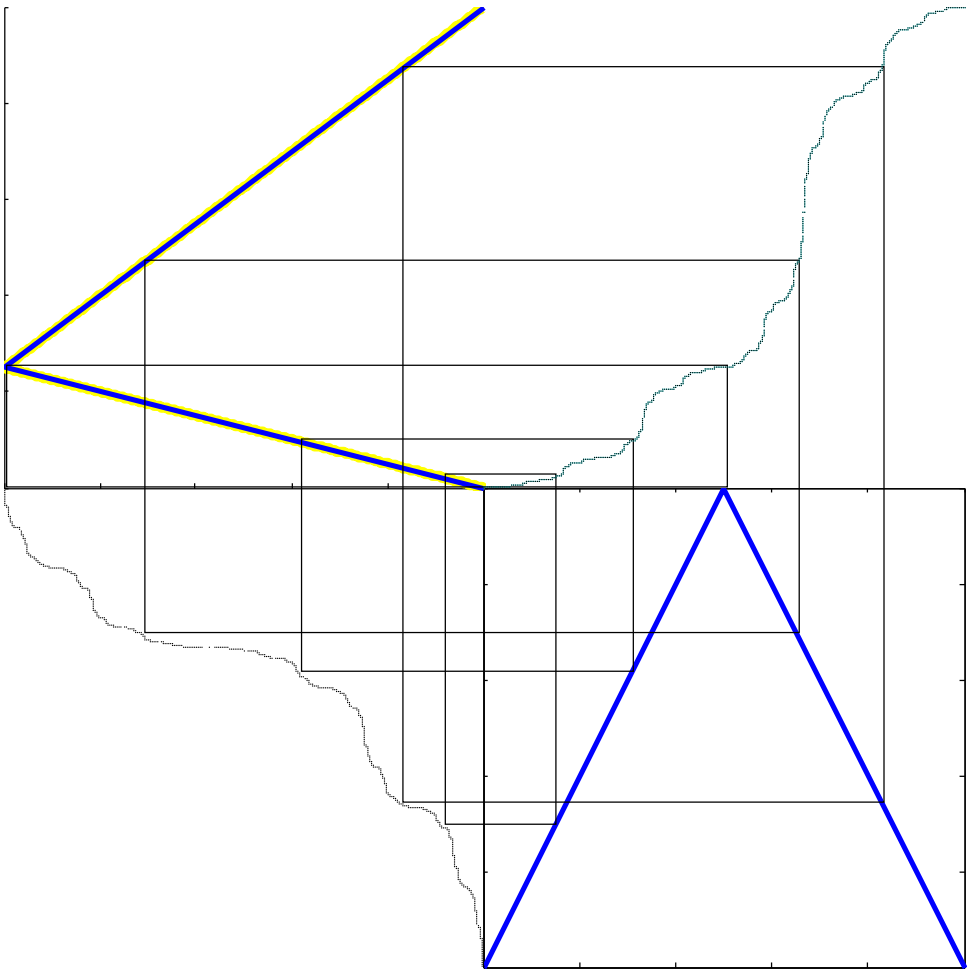


FIG. 1: **The conjugacy between two tent maps, shown as a quadweb.** Quadwebs allow us to visualize the commutative diagram, such as (1), with graphical data. The lower right panel shows a symmetric full shift tent map, F acting on space $X = [0, 1]$. The upper right panel shows a graph of the conjugacy function, h which maps $X = [0, 1]$ on the horizontal to $Y = [0, 1]$ on the vertical. The upper left panel shows skew tent map G acting on $Y = [0, 1]$, where we have oriented the graph by rotating the figure counterclockwise so that $y = h(x)$ lies in the domain of this graph. Similarly, the lower left is another copy of h, oriented to allow points x in the range of F to map to points y in the range of G. The black rectangles illustrate that the maps actually satisfy the commutative diagram. Interestingly, in this example, the conjugacy is strictly increasing, yet its derivative is 0 almost everywhere, as explained in Sec. 5. Quadwebbing is explained in more detail in Sec. 11, and therein, this particular quadweb is described further in Fig. 17.

comparing two dynamical systems becomes even more difficult in higher dimensions, but in this paper, we will focus only on dynamical systems embedded in the real line. As we will show in subsequent work, our methods do generalize in straightforward manner.

A primary interest of this work is to develop new methods to compare dynamical systems, whether they be equivalent in the sense of conjugacy, or not, and to do so in a manner which respects the notion of conjugacy. For two dynamical systems which are not conjugate, we will develop a new cost function to measure how

much the commuter function fails to be a conjugacy. Consider the following examples of indicative goal problems, which will be addressed after we have developed the appropriate tools in the early sections of this paper:

- **Goal 1**: We show that we can explicitly construct conjugacies. The two maps shown in Fig. 2(Left) are conjugate, even though we have not chosen the magic values, $r = 4$, and $a = 2$, for the logistic and tent maps respectively. For the first time, here we have developed methods to explicitly construct the conjugacy function, shown in Fig. 2(Right). Whereas the usual example, $r = 4$ and $a = 2$ gives a conjugacy function $h(x) = \frac{1}{2}[1 - \cos(\pi x)]$ which is actually a diffeomorphism, the more common situation for conjugate dynamical systems is as shown. The conjugacy function is merely continuous, and not differentiable. Because the Lyapunov exponents of g_r and T_a shown are different from each other (whereas diffeomorphisms preserve Lyapunov exponents) one would expect a discontininuity in the derivative of the conjugacy function $h(x)$ which relates the two. However, there is no clear one spot where the nondifferentiability should appear. In fact the resulting fractal-like function $h(x)$ is a Lebesgue singular function, as discussed in Sec. 5.

- **Goal 2** Here we will show that we can measure divergence from conjugacy, by measurements of the commuter function. We will quantify a concept of almost conjugacy in this paper. For example, by design, the full logistic map g_4 and the "noisy" logistic map $g(x)$ shown in Fig. 3(Left) are close under any of the usual $L^p([0,1])$ norms, as well as with respect to the sup-norm. It might be considered as a failing that primary method of comparison in the field of dynamical systems, the notion of conjugacy, is currently only able to give a boolean yes/no answer as to whether they are conjugate or not. It may seem obvious that the $g(x)$ shown is not conjugate to any one hump tent map. However, as may seem quite reasonable, the commuter function from $g(x)$ to the extended tent map $T_a(x)$ shown, does in fact appear to be at least $L^p([0,1])$ close to $h(x) = \frac{1}{2}[1 - \cos(\pi x)]$, which is the conjugacy between $g_4(x)$ and $T_2(x)$, as shown in Fig. 3(Right). The measure of mostly conjugate we develop in Sec. 7 measures the degree by which functions such as $f(x)$ fail each of the requirements of a homeomorphism.

In the pursuit of comparing two topologically dissimilar dynamical systems, one might consider relative entropies, whether they be differences of the topological entropies, or perhaps even the KS entropies, but these each give unsatisfactory answers in certain ways, some of which are outlined in Sec. 12. For example, considering the noisy logistic map in Fig. 3(Left), we realize that any semiconjugacy to a shift map requires a symbol space of many symbols, whereas the tent map shown requires only two symbols. While this suggests a topological dissimilarity between the two dynamical systems $g(x)$ and $T_a(x)$, it is not alone enough to cause large differences because the grammar of the symbolic dynamics of this noisy logistic map is highly restrictive. This is reflected by a submaximal topological entropy; the fact that there are 103 laps to the map suggests $h_T \leq \ln(103)$, and the fact that the slope of the map is a.a. 10 suggests $\geq \ln(10)$. However, in an experimental situation, depending on our ability to resolve the spatial fine scales of the thin humps, the map will appear more like the original logistic map of $h_T = \ln(2)$ which is much smaller than the maximal possibility. There is a notion of almost conjugacy already in the symbolic dynamics literature, [13], to compare two shift spaces, which relies on the existence of a factor map between the two shift spaces. Our work is complementary to that work, but it is somewhat different in that we can directly compare two maps which are not just shift maps and not just over shift spaces. Furthermore, we have the ability to explicitly construct the commuter function which is the realization of the factor map and was never before generally available even for the more restricted shift map setting. We also commend a related work by G.-C. Yuan *et. al.* [27] in which the problem of whether two observed time-series from the same physical chaotic process, (they suggest temperature and voltage from the same process as an example), can be identified as being dynamically related by conjugacy between one-d maps reconstructed from the data. They show that the cumulative distribution function constructed from the invariant measure μ from a dynamical system, $x_{n+1} = f(x_n), f : [c,d] \to [c,d]$, allows construction of a change of coordinates, $h : [c,d] \to [0,1]$, by, $w_n := h(x_n) := \mu([c,x_n])$. The new map, $g : [0,1] \to [0,1]$, is, $g(w_n) := h \circ f \circ h^{-1}(w_n)$. They call the map g a "canonical form" of f noting that it has uniform invariant density. Then they suggest that when the canonical forms of two maps f_1 and f_2 are the same, then the maps have been shown to be conjugate by construction. They further suggest a coordinate independent difference between two maps in terms of a relative entropy difference, $d(f_1, f_2) := d(g_1, g_2) := \int_0^1 [\log|g_1'(w)| - \log|g_2(w)|]dw$.

For now, we offer the statement as self evident that the noisy logistic $g(x)$ shown in Fig. 3(Left) and the logistic $g_4(x)$ are $L^p([0,1])$ close. However the KS entropy of the two are quite different: $h_{KS}(g) =$

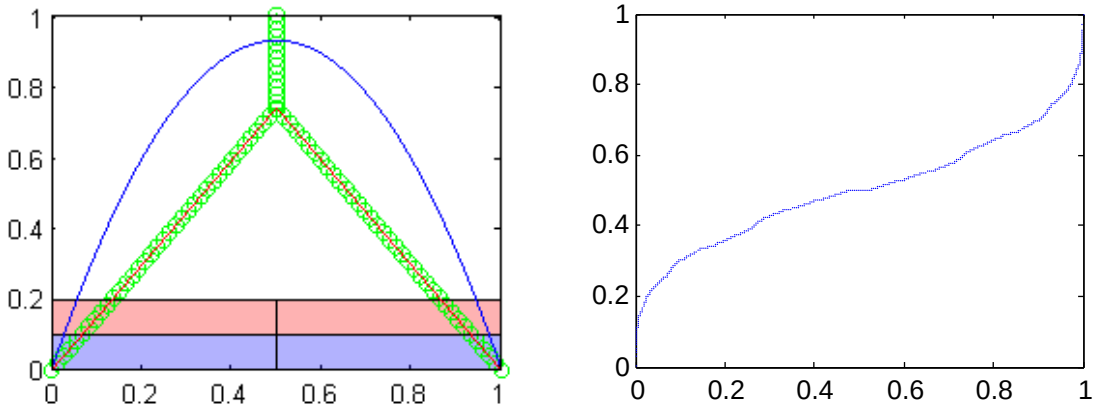


FIG. 2: (Left) A submaximal logistic map $g_r(x) = rx(1-x)$, $r < 4$ shown in blue is compared to a submaximal tent map, $T_a(y) = a(1-2|x-\frac{1}{2}|)$, $a < 2$, shown in green circles, together with a vertical extension which we call a *zed*, explained in Sec. 11 and Eq. (74). In this case, the regions matched are the same, as shown by the two horizontal bands at the bottom of the picture. (Right) The resulting conjugacy function $h(x)$ shown reminds of the better known $h(x) = \frac{1}{2}[1-\cos(\pi x)]$ which would result if we had chosen the famous values, $r = 4, a = 2$. However, the $h(x)$ shown is not smooth, and is apparently a Lebesgue singular function, as discussed in Sec. 5.

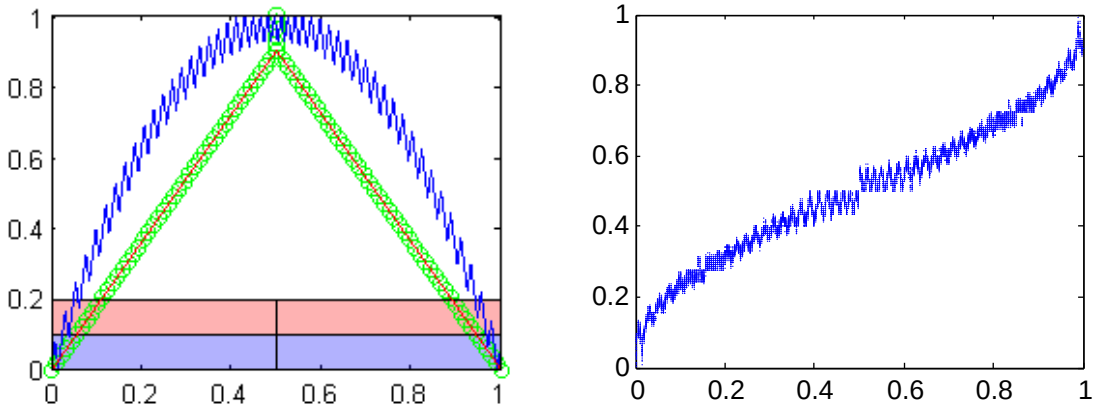


FIG. 3: (Left) A "noisy" logistic map g. This map is obviously nearby the full logistic map g_4, in both any $L^p([0,1])$, $p \geq 1$, as well as in the sup-norm, but it is not in the $C^1([0,1])$ norm as this map g has a slope $|g'(x)| = 10$ everywhere which it exists. Also shown in green circles is a tent map with vertical extension which gives the "best" almost conjugacy within the family of symmetric tent maps $T_a(x)$. (Right) The resulting commuter function f between g and this tent map is reminiscent of the homeomorphism $h(x) = \frac{1}{2}[1-\cos(\pi x)]$ between the full logistic map and the full tent map. However, where h is a diffeomorphism, f shown is not even a homeomorphism, but it is almost a conjucay in a manner to be defined in Sec. 7; in this sense, we will call the noisy logistic map g as *almost conjugate* to the tent map T_a shown.

$\ln(10) > h_{KS}(g_4) = \ln(2)$. [We remark that since $h_{KS}(g) = \ln(m)$, where m is the slope of the sawtoothing, with $m = g'(x)$, whereever the derivative exists. Consequently, by increasing the slope of the sawtooth, we can make the entropy difference arbitrarily large while keeping g and g_4 "close" in the $L^p([0,1])$ topology.] Clearly, KS entropy *may* be a poor measure to compare two maps which seem to be dynamically close. Likewise, topological entropy h_T can be made arbitrarily largely different between a sawtoothed g which is L^p-close to g_4. By directly measuring the "best" commuter function f, we have designed a new and useful way to compare similarity between dynamical systems.

2. BACKROUND

Here we explicitly recall well known notions in the field of dynamical systems for reference within our writings to follow, because a major aspect of this paper will include discussion of how each of the parts of the definitions may fail when attempting to compare two dissimilar dynamical systems. Stated for discrete time-mappings, and as found in many excellent texts such as [6, 11],

Definition: Given a dynamical systems, $g_1 : X \to X$, in a topological space, $(X, \mathcal{A})$, and a second dynamical system, $g_2 : Y \to Y$, in a topological space $(Y, \mathcal{B})$, these mappings are topologically *conjugate* if there exists a mapping $h : X \to Y$, such that

- h is a homeomorphism between the spaces X and Y (thus the spaces are topologically equivalent),
- h commutes with respect to the two mappings: $h \circ g_1(x) = g_2 \circ h(x)$ for all $x \in X$.

The second condition implies the commuting diagram,

$$\begin{array}{ccc} X & \xrightarrow{g_1} & X \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ Y & \xrightarrow{g_2} & Y \end{array} \tag{1}$$

allows iteration of the map on the top, $x \mapsto x' = g_1(x)$, followed by the change of coordinates $y' = h(x')$ gives the same result as first applying the change of coordinates to the initial condition $y = h(x)$, followed by the lower mapping, $y \mapsto y' = g_2(y)$.

Homeomorphism is the standard of equivalence between two topological spaces [14],

Definition Two topological spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are *homeomorphic* if there is a mapping between the two spaces, $h : X \to Y$, such that h is,

1. one-one,
2. onto,
3. continuous,
4. the inverse function, h^{-1} exists and it is continuous.

The function h satisfying this definition is called a *homeomorphism.*

The commuting homeomorphism function h in the definitions above is called a *conjugacy.* In subsequent sections, we will put forward algorithms to construct h when it exists, by demanding the commuting property. When the conjugacy does not exist, we still have a reasonable way of forcing the commuting diagram to hold, by a mapping which we call a *commuter* (now denoted as f), between X and Y . We examine the commuter f to measure the manner in which each and any of the above requirements for a homeomorphism may fail.

3. CONJUGACY OF LINEAR MAPS

Consider two linear maps M_1 and M_2 defined by

$$\begin{aligned} M_1(x) &:= m_1 x \\ M_2(x) &:= m_2 x. \end{aligned} \tag{2}$$

Topological conjugacy is the hallmark of equivalence in the field of dynamical systems. So it is reasonable to ask, "given two linear maps, under what conditions are the dynamical systems conjugate to each other?"

If they are conjugate, then we ought to be able to find the conjugacy h between the two maps, where h is a homeomorphism satisfying the commutative diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{M_1} & \mathbb{R} \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ \mathbb{R} & \xrightarrow{M_2} & \mathbb{R}. \end{array} \tag{3}$$

Then h must satisfy

$$h \circ M_1 = M_2 \circ h, \tag{4}$$

or (more directly)

$$h(m_1 x) = m_2(h(x)). \tag{5}$$

As a simple example, we note that if $m_1 = 2$ and $m_2 = 4$, then we require

$$h(2x) = 4h(x),$$

which is satisfied (for positive x) by $h(y) = y^2$. Motivated by this example, we search for a conjugacy formula by assuming the ansatz

$$h(x) = c\,\mathrm{sign}(x)|x|^\gamma, \qquad c \neq 0, \quad \gamma > 0, \tag{6}$$

where the slight modifications to a simple power law are required to ensure that h is 1-1 and onto the real line. We seek to identify c and γ to satisfy (4). Substitition of this ansatz into (5) yields

$$c\,\mathrm{sign}(m_1 x)|m_1|^\gamma |x|^\gamma = m_2 c\,\mathrm{sign}(x)|x|^\gamma. \tag{7}$$

We immediately find that $c \neq 0$ is arbitrary. Additionally, since this equation should apply when $x \neq 0$, we may divide by $|x|^\gamma$:

$$\mathrm{sign}(m_1 x)|m_1|^\gamma = m_2\,\mathrm{sign}(x). \tag{8}$$

By applying the *signum* function to both sides, we can conclude

$$\mathrm{sign}(m_1) = \mathrm{sign}(m_2) \tag{9}$$

will be required, which is consistent with M_1 being conjugate to M_2. By solving (8) for γ, we find

$$\gamma = \frac{\log|m_2|}{\log|m_1|}. \tag{10}$$

Since γ must be positive, we infer that either (a) $m_1 < 1$ and $m_2 < 1$, or (b) $m_1 > 1$ and $m_2 > 1$. Again, this requirement is consistent with the two maps being conjugate, since they must either both be contracting or both be expanding. Because these systems have a single periodic point, there is significant flexibility in describing a conjugacy, and we have simply identified one choice. (For an alternative treatment, see Robinson [11].) It is interesting to note that for this simple linear example, the action of the conjugacy function on the state space variable (x) can be interpreted as a time dilation.

4. FIRST CONSTRUCTION OF COMMUTERS: GOAL 1

In this section, we will take the tent map as an easy setting to introduce a centerpiece of this paper, which is how a the commuter, or would be conjugacy, can be explicitly constructed by a fixed point iteration scheme. In this simple setting, we will rigorously prove several aspects of existence, and convergence of the fixed point iteration. In subsequent sections, we will see that for more complicated problems, the methods seem to work numerically well beyond our sufficient theorems.

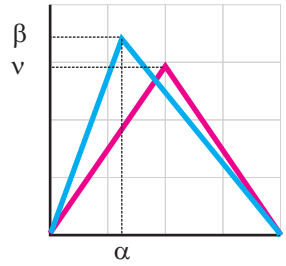
4.1. A conjugacy between tent maps

Consider the family of (skew) tent maps $s(x)$ defined on $[0,1]$, with the following restrictions:

- $s(0) = 0$, and $s(1) = 0$.
- The peak of the tent occurs at $s(\alpha) = \beta$, with $0 < \alpha < 1$.
- To ensure that the map is locally expanding, we require.

$$\max(\alpha, 1-\alpha) < \beta \leq 1, \tag{11}$$

We call this family $\mathcal{S}$, where the family is parameterized by α and β, the coordinates of the peak of the tent. We denote a specific member of this family as $S_{\alpha,\beta}$, where the coordinates of the peak of that tent are at (α, β). Within this family of skew tents, we consider that subset of maps that are symmetric about $x = 1/2$. We denote this sub-family as $\mathcal{T}$, where $\mathcal{T} \subset \mathcal{S}$. An arbitrary member $T_\nu \in \mathcal{T}$ is defined by $T_\nu := S_{1/2,\nu}$.



First, we note the following lemma regarding the existence of a conjugacy:

Lemma 1. *Let $S_{a,b}$ be a particular member of $\mathcal{S}$. Then there exists a ν_0 such that $S_{a,b}$ is conjugate to $T_{\nu_0} \in \mathcal{T}$.*

Sketch of Proof: Kneading sequences need to be matched, and we know that there is a value of ν_0 which matches the kneading sequences and hence symbolic dynamics by a so-called intermediate value theorem of kneading sequences, which can be found in either Misiurewicz and Visinescu [23] or Collet and Eckmann[12]. Although matching symbolic dynamics is not enough to prove that the maps are conjugate, it is necessary. Similarly the fact that both maps are strictly monotone is enough to guarantee that monotone laps map to monotone laps, from which it can be shown that there is a match between any two points with the same symbolic itineraries. The match of symbolic dynamics does guarantee a match between the eventually periodic points. Because both maps are are assumed to be everywhere expanding, the matching can be extended to the entire interval, with h aand h^{-1} continuous. □

The remainder of this section is constructed under the presumption that we have chosen a particular value for a and b, and though they are arbitrary, they remain fixed. The map $S \equiv S_{a,b}$ will be used to symbolize this arbitrary but fixed map. To find the conjugacy whose existence is indicated by Lemma 1, we need to solve the functional equation

$$S \circ h(x) = h \circ T(x). \tag{12}$$

In particular, one must find an h that not only satisfies (12) but is also a homeomorphism. In general, there is no direct technique to find such an h. Instead, we propose an indirect solution approach: we create a fixed point iteration scheme that converges to a solution. The rest of this section describes the development of that iterative scheme, which we decompose into three components:

- Creating a contraction mapping which generates solutions to the commutative diagram.
- Explaining why the result of that contraction yields a conjugacy h when S and T are conjugate (or, equivalently, when $\nu = \nu_0$, with ν_0 known).
- Describing an iterative technique to find ν_0 when a and b are given, but the required conjugacy parameter value (from Lemma 1) is not known.

4.2. A contraction mapping from the commutative relationship.

On the interval $[0, 1/2]$, the explicit form for the symmetric tent map is $T(x) = 2\nu_0 x$. Substituting into (12) gives

$$S \circ h(x) = h(2\nu_0 x). \tag{13}$$

To write an explicit description of S, we note that the interval $[0, 1/2]$ describes the domain of the left part of the tent map T. Because we are identifying the *conjugacy* between these two maps (which preserves the kneading sequence), we conclude that $h[0, 1/2] = [0, a]$ because $[0, a]$ is the domain of the left part of the S map. On that interval, $S(u) = \frac{b}{a}u$, and substitution into (13) gives

$$\frac{b}{a}h(x) = h(2\nu_0 x). \tag{14}$$

Using a similar logic, but applied to the interval $(1/2, 1]$ gives

$$\frac{b}{1-a}(1 - h(x)) = h(2\nu_0(1 - x)). \tag{15}$$

Therefore, the conjugacy function $h(x)$ must satisfy the functional equation,

$$h(x) = \begin{cases} \dfrac{a}{b}h(2\nu_0 x) & 0 \le x \le 1/2 \\ 1 - \dfrac{1-a}{b}h\left(2\nu_0\left(1 - x\right)\right) & 1/2 < x \le 1. \end{cases} \tag{16}$$

(See Fig 4 for a graphical representation of this relationship.)

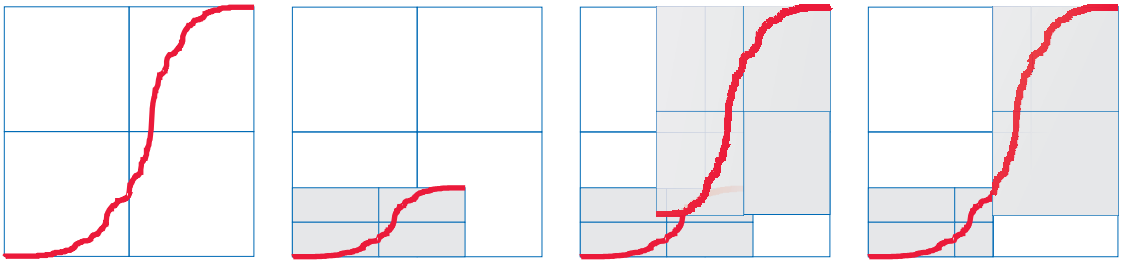


FIG. 4: **Satisfying the functional equation.** (16) can be viewed as a process: take $h(x)$ (panel 1) and make a copy, shrunken by a/b in the vertical and by $2\nu_0$ in the horizontal (panel 2). Take a second copy, scaled the same horizontally, but vertically scaled by $(1 - a)/b$. Rotate this copy by 180 degrees and place in the upper right portion of the unit square (panel 3). Then truncate the left copy to the interval $[0, 1/2)$ and the right copy to $[1/2, 1]$. The result (panel 4) should return the original $h(x)$.

As an additional comment resulting from analysis of the kneading sequences, we note that

$$h(1/2) = a. \tag{17}$$

Additionally, evaluating (16) at $x = 1/2$, we have

$$h(1/2) = a = \frac{a}{b}h(\nu_0),$$

yielding that

$$h(\nu_0) = b. \tag{18}$$

Using (16) as a guide, we now create an operator whose fixed point will satisfy the commutative diagram. We consider the space $B([0, 1], \mathbb{R})$, the set of all bounded functions from $[0, 1]$ to the real numbers, which is a Banach space, with norm given by

$$\|f\| \equiv \|f\|_\infty := \sup_{x\in[0,1]} |f(x)|.$$

Then we form the closed subset $\mathcal{F} \subset B([0,1],\mathbb{R})$, defined by

$$\mathcal{F} = \{f | f : [0,1] \to [0,1]\}, \tag{19}$$

the set of functions from $[0,1]$ to $[0,1]$. Then given (a,b) satisfying $\max(a, 1-a) < b < 1$, we define a one parameter family of operators $M_\nu : \mathcal{F} \to \mathcal{F}$ for $1/2 < \nu \leq 1$:

$$M_\nu f(x) := \begin{cases} \dfrac{a}{b} f(2\nu x) & 0 \leq x \leq 1/2 \\ 1 - \dfrac{1-a}{b} f\left(2\nu\left(1-x\right)\right) & 1/2 < x \leq 1. \end{cases} \tag{20}$$

The constraints on the parameters a, b and ν are required to ensure that $\mathcal{F}$ is mapped into itself, but they also cause the operator to be a contraction.

Lemma 2. *M_ν is a uniform contraction on $\mathcal{F}$, where the contraction is with respect to $\|\cdot\|_\infty$.*

Proof. Define

$$\lambda = \max\left(\frac{a}{b}, \frac{1-a}{b}\right).$$

Then $0 < \lambda < 1$. We compute

$$\|M_\nu f_1 - M_\nu f_2\|_\infty = \sup_x |M_\nu f_1(x) - M_\nu f_2(x)|.$$

We decompose this problem into the two cases $x \leq 1/2$ and $x > 1/2$. For the first case,

$$\begin{aligned} \sup_{0\leq x\leq 1/2} |M_\nu f_1(x) - M_\nu f_2(x)| &= \sup_{0\leq x\leq 1/2} \left|\frac{a}{b}\left(f_1(2\nu x) - f_2(2\nu x)\right)\right| \\ &\leq \frac{a}{b} \sup_{0\leq y\leq 1} |f_1(y) - f_2(y)| \leq \lambda \|f_1 - f_2\|. \end{aligned}$$

Similarly

$$\begin{aligned} \sup_{1/2<x\leq 1} |M_\nu f_1(x) - M_\nu f_2(x)| &= \sup_{1/2<x\leq 1} \left|\frac{1-a}{b}\left(f_1(2\nu x) - f_2(2\nu x)\right)\right| \\ &\leq \frac{1-a}{b} \sup_{0\leq y\leq 1} |f_1(y) - f_2(y)| \leq \lambda \|f_1 - f_2\|. \end{aligned}$$

So M_ν is a contraction, with contraction constant λ. Because λ does not depend on ν, the contraction is *uniform.* □

With this groundwork in place, we establish the existence of a fixed point of the operator.

Lemma 3. *There is a unique $f_\nu \in \mathcal{F}$ satisfying*

$$M_\nu f_\nu = f_\nu. \tag{21}$$

Moreover, for an arbitrary $f_0 \in \mathcal{F}$, if we define the sequence of functions

$$f_{n+1} = M_\nu f_n,$$

this sequence will converge to f_ν :

$$f_\nu := \lim_{n\to\infty} f_n. \tag{22}$$

Proof. The lemma is a direct application of the Banach-Caccioppoli Contraction Mapping Principle [15], and we simply need to verify that we satisfy the hypothesis of the theorem. Since $B([0,1],\mathbb{R})$, is a Banach space, with $\mathcal{F}$ a closed subset, and $M_\nu : \mathcal{F} \to \mathcal{F}$, the theorem applies, and the conclusions are immediate. □

Remark: For any chosen ν, by construction, the fixed point f_ν will satisfy the requirements of the commutative diagram, at least over some portion of its domain (See Appendix 11 for more detail). The question remains as to whether it is a conjugacy function between the two maps S and T_ν. If the two maps are conjugate, which we know is possible by Lemma 1 by choosing the parameter to be $\nu = \nu_0$, then it is straightforward to check that the f_{ν_0} produced by the fixed point iteration scheme is that conjugacy function, and we call it $h(x) = f_{\nu_0}(x)$. It is important to note for much of the following development, that even if the two maps S and T_ν are not conjugate and so no conjugacy function $h(x)$ exists, the fixed point iteration still produces a function f_ν satisfying the communative diagram. However, it cannot be a homeomorphism.

4.3. The Search For ν_0, A Problem of Parameter Estimation

Suppose we are given a chaotic map from the family $\mathcal{S}$. (In other words, (a, b) are fixed.) By Lemma 1, we know that there is a ν_0 such that map T_{ν_0} is conjugate to S. In this section, we give a short explanation of how we find that conjugate map.

When $\nu \neq \nu_0$, the maps $S_{a,b}$ and T_ν are not conjugate, and while the existence of f_ν as the fixed point of a contraction is guaranteed by Lemma 3, f_ν fails to be a homeomorphism. We find that for $\nu < \nu_0$, f_ν is monotone increasing, but not continuous and not onto $[0, 1]$. The largest "gap" is at $x = 1/2$. Because f_ν satisfies a recursive formulation (which creates a self-similar structure), a scaled version of this defect appears at all x coordinates associated with the kneading sequence of T. We define

$$\tilde{\lambda}(\nu) := \lim_{x \to (1/2)^-} f_\nu(x) - a \tag{23}$$

as a way of measuring this "defect," with $\tilde{\lambda}(\nu) < 0$ when $\nu < \nu_0$. (See Fig 5.) We discuss measuring homeomorphism defects and easier to compute suitable surrogates in Sec. 8.

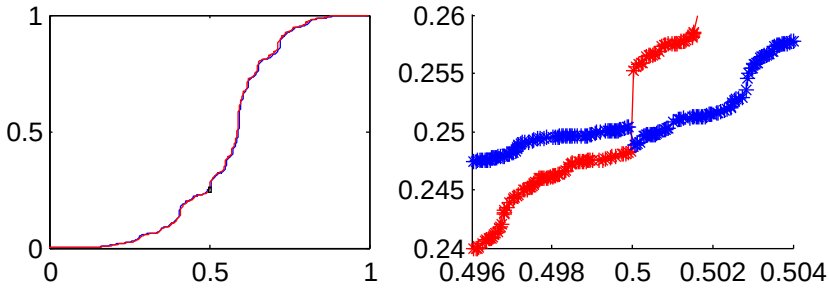


FIG. 5: **Resolving** ν_0. Fixing $a = 0.25$ and $b = 0.9$, we graph f_ν for $vu = 0.72$ (red) and $\nu = 0.73$ (blue). The zoomed panel (right) illustrates that $\tilde{\lambda}(0.72) < 0$ (the red curve jumps up at $x = 0.5$) while $\tilde{\lambda}(0.73) > 0$ (the blude curve jumps down).

As a short explanation for the gaps, we note that since

$$h_T(T_\nu) < h_T(T_{\nu_0}) = h_T(S_{a,b}),$$

the only way that we match equivalent orbits under the T and S dynamics is to restrict the domain of S. In other words, there must be orbits of S that cannot be represented by T, and the "gaps" indicate those initial conditions in S whose trajectories do not have a matching trajectory under T.

In the case $\nu > \nu_0$, the entropy mismatch is in the other direction, with

$$h_T(T_\nu) > h_T(T_{\nu_0}) = h_T(S).$$

There are now orbits in T that cannot be represented in S. Consequently, the commuting function f_ν must "remove" an appropriate amount of entropy from the T dynamics by mapping multiple points in the domain to single points in the range, so f_ν is not 1-1. Numerically, we observe similar discontinuities at $x = 1/2$ (repeated in accordance with the kneading sequence). We use the same measuring function $\tilde{\lambda}(\nu)$ as defined in (23), but we note that $\tilde{\lambda}(\nu) > 0$ when $\nu > \nu_0$.

Despite any limitations of numerical approximations of $\tilde{\lambda}(\nu)$, we know analytically that

$$\tilde{\lambda}(\nu_0) = 0.$$

Consequently, we can apply typical scalar root finding algorithms (bisection, or secant algorithms, for example) to approximate ν_0.

5. RELATIONSHIP TO DE RHAM FUNCTIONS AND LEBESGUE SINGULAR FUNCTIONS

In this section, we will show some interesting properties of typical homeomorphisms between full shift tent maps. We show that the conjugacy function $h(x)$ which results is both a Lebesgue singular function, and in many ways reminds us of a de Rham function. The similarity of our homeomorphisms $h(x)$ to the de Rham functions arises from the fact that both are solutions to functional equations of similar form. We recall definitions of these peculiar functions before we proceed to prove the properties of our $h(x)$. While we have not proven that homeomorphisms between two tents which are not both full have these properties, and we also do not prove these properties for either non-tent maps or non-full shift tent maps, the properties are certainly suggestive of the nature of the kinds of peculiarities one might expect, and even which seem apparent in the simulations. Most striking is the degree to which we now know that the usual homeomorphism example between a full tent and the full logistic map, which gives rise to a diffeomorphism, can mislead intuition from the more typical conjugacy.

5.1. Lebesgue Singular Functions

Definition [16]: A continuous function of bounded variation $s(x)$ is called a *singular function* if it is differentiable almost everywhere on its domain, and the derivative $s'(x) = 0$ where it exists, and it is called a *Lebesgue singular function* when almost every is in the sense of Lebesgue measure.

Definition [18, 19]: *de Rham functions* A de Rham function is the unique bounded solution of the functional equation

$$\begin{aligned} \phi(\frac{t}{2}) &= a\phi(t) \\ \phi(\frac{t+1}{2}) &= a + (1-a)\phi(t), \end{aligned} \tag{24}$$

where $a \in (0,1)$ is fixed and (24) is satisfied for all $t \in [0,1]$.

5.2. $h(x)$ is a Singular Function.

In this section we will show that for some parameter values, the f_ν which results from Lemma 3 is a singular function. The argument of this proof follows closely the structure found in [24], which is standard for the De Rham functions. However, slight modifications to the proof are required due to the differences between his problem and ours. We will specifically consider the conjugacy between a full shift skew tent map and the full shift symmetric tent.

Consider the set of skew tent maps of the form $S_{a,1}$, which are the maps which generate a full shift on two symbols. These maps are conjugate to the full shift symmetric map, T_1, by which we infer that $\nu = \nu_0 = 1$. If $a = 1/2$, then $S \equiv T_1$, and the conjugacy would be the identity map. However, when $a \neq 1/2$, we can find the conjugacy h by developing the contraction operator and identifying its fixed point. Then $f_\nu \equiv f_{\nu_0} \equiv f_1 \equiv h$. Because h is a conjugacy, we know it is strictly increasing and continuous. However, we will find that despite these restrictions, h remains a very strange function.

Recall [16, 17] that by the *Lebesgue Decomposition Theorem*, a function $g(x)$ of Bounded variation may be decomposed

$$g(x) \equiv a(x) + s(x), \tag{25}$$

where $a(x)$ is absolutely continuous and $s(x)$ is purely singular. As additional description, of this decomposition, we have

$$a(x) = \int g'(x)\mathrm{d}x \Longrightarrow g(x) = s(x) + \int g'(x)\mathrm{d}x.$$

So the singular part separates out the components that prevent a function from satisfying the fundamental theorem of calculus. Because h is strictly monotone and continuous on a closed interval, it is of bounded variation, so it is decomposeable as above. However, in the following paragraphs, we show that $h(x)$ is itself a purely singular function, such that

$$h(x) \equiv a(x) + s(x) \Longrightarrow a(x) \equiv 0.$$

Proposition 1. *Let $h(x)$ be the conjugacy between $S_{a,1} \in \mathcal{S}$ and T_1, where $a \neq 1/2$. Then $h'(x)$ exists for almost every $x \in [0,1]$. Moreover, $h'(x) = 0$ wherever $h'(x)$ exists.*

Proof. Because h must be a homeomorphism, it is increasing. Then a standard result from analysis [16] tells us that h is a.e. differentiable, giving the existence of $h'(x)$. Suppose that the derivative exists at some $0 < x < 1$; we will show that $h'(x) = 0$.

For each n, we find integer k_n such that

$$x \in I_n := \left(\frac{k_n}{2^n}, \frac{k_n+1}{2^n}\right];$$

measuring the range of h over the interval I_n, we define

$$p_n := h\left(\frac{k_n+1}{2^n}\right) - h\left(\frac{k_n}{2^n}\right). \tag{26}$$

Since $h'(x)$ exists, we know that

$$h[(k_n)2^{-n}, (k_n+1)2^{-n}] = \frac{p_n}{2^{-n}} \to h'(x),$$

where the left hand side symbolizes the Newton divided difference. If $h'(x) \neq 0$, then the ratio of two successive divided differences (for n and $n+1$) will tend to 1. Therefore the ratio

$$\frac{p_{n+1}}{p_n} \to \frac{1}{2}, \tag{27}$$

as long as $h'(x)$ is distinct from 0.

To complete the proof, we will show that either $\frac{p_{n+1}}{p_n} = a$ or $\frac{p_{n+1}}{p_n} = 1-a$, where this property holds for all x. This portion of proof results from the self-similar structure which derives from the fact that $h(x)$ is the fixed point of the contraction mapping M_1 defined by (20), where we are using $b = \nu = \nu_0 = 1$. The proof procedes by induction:

- **True for** $n = 0$. We have $p_0 = h(1) - h(0) = 1 - 0 = 1$. If $x \leq 1/2$, then $I_1 = (0, 1/2]$; $h(1/2) = a$, and $\frac{p_1}{p_0} = a$. For $x > 1/2$, $p_1 = h(1) - h(1/2)$, giving $\frac{p_1}{p_0} = 1 - a$.
- **Induction.** Assume that for arbitrary x, $\frac{p_{j+1}}{p_j} \in \{a, 1-a\}$ for all $0 \leq j \leq n-1$. We will show that the property holds for $j = n$:
  - Note that the endpoints of I_n are $x_{left} := \frac{k_n}{2^n}$ and $x_{right} := \frac{k_n+1}{2^n}$. Denote the midpoint of I_n as x_{mid}. Then either $I_{n+1} = (x_{left}, x_{mid}]$ (which implies that $k_{n+1} = 2k_n$) or $I_{n+1} = (x_{mid}, x_{right})$, implying $k_{n+1} = 2k_n + 2$. If we assume that I_{n+1} is on the left half of I_n, then

$$\frac{p_{n+1}}{p_n} = \frac{h\left(\frac{2k_n+1}{2^{n+1}}\right) - h\left(\frac{2k_n}{2^{n+1}}\right)}{h\left(\frac{k_n+1}{2^n}\right) - h\left(\frac{k_n}{2^n}\right)}, \tag{28}$$

whereas if I_{n+1} is the right half of I_n, we have

$$\frac{p_{n+1}}{p_n} = \frac{h\left(\frac{2k_n+2}{2^{n+1}}\right) - h\left(\frac{2k_n+1}{2^{n+1}}\right)}{h\left(\frac{k_n+1}{2^n}\right) - h\left(\frac{k_n}{2^n}\right)}, \tag{29}$$

We now take advantage of the self-similarity implied by (16): When $x \leq 1/2$, we have that $ah(2x) = h(x)$. So for each function evaluation in (28) and (29), we perform the substitution

$$ah\left(\frac{m}{2^{r-1}}\right) = h\left(\frac{m}{2^r}\right). \tag{30}$$

To simplify notation, we denote by I'_n the dyadic intervals containing the point $2x$, and define p'_n as the length of the intervals $h[I'_n]$. Then applying (30) to (28) and (29), we have that

$$\frac{p_{n+1}}{p_n} = \frac{ah\left(\frac{2k_n+1}{2^n}\right) - ah\left(\frac{2k_n}{2^n}\right)}{ah\left(\frac{k_n+1}{2^{n-1}}\right) - ah\left(\frac{k_n}{2^{n-1}}\right)} = \frac{p'_n}{p'_{n-1}}, \tag{31}$$

or

$$\frac{p_{n+1}}{p_n} = \frac{ah\left(\frac{2k_n+2}{2^n}\right) - ah\left(\frac{2k_n+1}{2^n}\right)}{ah\left(\frac{k_n+1}{2^{n-1}}\right) - ah\left(\frac{k_n}{2^{n-1}}\right)} = \frac{p'_n}{p'_{n-1}}. \tag{32}$$

By assumption,

$$\frac{p'_n}{p'_{n-1}} \in \{a, 1-a\}.$$

The argument for $x > 1/2$ is similar, though the scaling factor is $1-a$ instead of a.

□

5.3. Remarks and Lessons from the de Rham-like Conjugacies

From Sec. 3, we saw a diffeomorphism between two linear maps of the form, Eq. (6). We have remarked that since two piecewise linear maps with different metric entropies must not be diffeomorphically related, following Eqs. (77)-(78), then any conjugacy between two such maps cannot be everywhere differentiable. It may not immediately obvious what is the distinguished point or points where the discontinuity of the derivative of $h(x)$ should reside. Now, in light of the analysis of the de Rham like properties of $h(x)$ in the previous section, we see that the discontinuities are dense, located at endpoints of the dyadic intervals I'_n, corresponding to preimages of the map peaks, where there is a discontinuity in the slope of the nth composition of the symmetric tent map.

6. CONSTRUCTING THE CONTRACTION MAPPING FOR GENERAL TRANSFORMATIONS

In this section, we consider more general classes of dynamical systems than the tent maps examined in §4 4.1. Our goal is to develop a methodology that will allow us to "solve" the commutative diagram. Our primary tool will be the construction of an appropriate operator whose fixed point yields a solution to the commutative diagram. We call this operator the *commutation operator*. Our approach to defining the operator will be to generalize the technique of §4 4.1 which constructed a contraction operator to solve the appropriate functional equation. We explain our method in two parts, first by showing the construction process when the dynamical systems are conjugate, and then we show how to adapt that method to the case when the maps are not conjugate. For this general case, we do not prove that the *commutation operator* is a contraction, but on all numerical examples we have tested, it has converged to an approximate solution.

6.1. Constructing the *commutation operator* for conjugate maps.

Consider the diagram

$$\begin{array}{ccc} X & \xrightarrow{g_1} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow{g_2} & Y, \end{array} \tag{33}$$

where g_1 and g_2 are 1-D chaotic maps. The commutative diagram is equivalent to saying that f is a solution to the functional equation

$$f \circ g_1 = g_2 \circ f. \tag{34}$$

To phrase this equation as a fixed point problem, we would like to solve for f on the right hand side, but because g_2 is chaotic, it will not be invertible. As we did with the tent maps, we tackle this problem by providing a piecewise definition for the operator, where on each piece we use a restricted subset of Y to define the inverse of g_2.

In this subsection, we assume that g_1 and g_2 are conjugate. Therefore there is a homeomorphism h which satisfies the commutative diagram, and this h is called a conjugacy. It is our goal to have h be a fixed point of our constructed operator. We demand that h must map monotone segments of the graph of g_1 onto monotone segments of g_2. To simplify the explanation, we assume that X and Y are compact intervals. Let $\boldsymbol{x} = \{x_0, x_1, \ldots, x_n\}$, with $X \equiv [x_0, x_n]$, and g_1 monotone on each interval $[x_{i-1}, x_i]$, for $i = 1, \ldots, n$. This notation results in a natural decomposition of X into a union of disjoint subintervals,

$$I_{X1} = [x_0, x_1], \tag{35}$$
$$I_{Xi} = (x_{i-1}, x_i], \qquad i = 2, \ldots, n, \tag{36}$$
$$X = \bigcup_i I_{Xi}, \tag{37}$$

where we have (arbitrarily) chosen the closure conditions of the intervals. Because of the conjugacy of the dynamical systems, there must exist $\boldsymbol{y} = \{y_0, \ldots, y_n\}$, with $Y \equiv [y_0, y_n]$ and g_2 monotone on each interval $[y_{i-1}, y_i]$, with a similar decomposition of Y into a union of disjoint subintervals I_{Yi}. [28] To allow for well defined inverse functions, we denote

$$g_{2i} = g_2 | I_{Yi}.$$

Because f is a conjugacy, it must map monotone intervals of X to monotone intervals of Y, such that we may write

$$f[I_{Xi}] = I_{Yi}, \qquad i = 1, \ldots, n, \tag{38}$$

by which we infer the identity

$$f[I_{Xi}] = g_{2i}^{-1} \circ g_2 \circ f[I_{Xi}]. \tag{39}$$

We can now rewrite (34) as a system of equations,

$$g_{2i}^{-1} \circ f \circ g_1[I_{Xi}] = f[I_{Xi}], \qquad i = 1, \ldots, n. \tag{40}$$

Having now solved for f on the RHS of (40), we can use this formulation to construct the commutation operator. Let $\mathcal{F}$ be the set of functions from X to Y. Then we define the operator $\mathfrak{C}_{g_1}^{g_2} : \mathcal{F} \to \mathcal{F}$, which takes $F \in \mathcal{F}$ to $\mathfrak{C}_{g_1}^{g_2} F$ by

$$\mathfrak{C}_{g_1}^{g_2} F(x) = \hat{g}_{2i}^{-1} \circ F \circ g_1(x) \qquad x \in I_{Xi}. \tag{41}$$

Note the slight change in notation between (40) and (41) in that we now use $\hat{g}_{2i}^{-1}$ instead of simply g_{2i}^{-1}, where we still need to define and explain the former. We remark that g_{2i}^{-1} is defined only for the interval $g_2[I_{Yi}]$, which may not be all of Y. However, $F \circ g_1(x)$ may be any value in Y. Consequently, we need to extend g_{2i}^{-1} to all of Y. We define $\hat{g}_{2i}^{-1}$ such that it satisfies:

- $\hat{g}_{2i}^{-1}$ is continuous on Y,
- $\hat{g}_{2i}^{-1} \equiv g_{2i}^{-1}$ on $g_2[I_{Yi}]$,
- $\hat{g}_{2i}^{-1}$ is Lipshitz continuous on $Y - g_2[I_{Yi}]$, with Lipshitz constant $L < 1$.

Simply by expanding the domain of definition for each of the inverses, we ensure that for each $x \in X$, $\mathfrak{C}_{g_1}^{g_2} F(x)$ exists, so that the commutation operator is well defined. The requirement on the Lipschitz constant is meant to increase the likelihood that the operator may be a contraction. However, the requirements, as stated, do not uniquely define $\hat{g}_{2i}^{-1}$. We have chosen to retain this flexibility because we find that a judicious choice of extension may ease the numerical implementation and provide additional insight into the relationship between the systems. In this paper, the examples have used one or the other of the following strategies: (1) assume $\hat{g}_{2i}^{-1}$ is constant on $Y - g_2[I_{Yi}]$, or (2) when g_2 is piecewise linear, then simply assume that $\hat{g}_{2i}^{-1}$ lies on the same line as g_{2i}^{-1}.

Proposition 2. *Let g_1 and g_2 be conjugate, with $h : X \to Y$ the conjugacy. Then h is a fixed point of the* commutation operator, $\mathfrak{C}_{g_1}^{g_2}$.

Proof. Take an arbitrary x. Then there is a i such that $x \in I_{Xi}$. Because h is a conjugacy, it must map monotone intervals to monotone intervals. In particular, $h[I_{Xi}] = I_{Yi}$, which implies that $y := h(x) \in I_{Yi}$. We apply the definition of the commutation operator,

$$\mathfrak{C}_{g_1}^{g_2} h(x) = \hat{g}_{2i}^{-1} \circ h \circ g_1(x). \tag{42}$$

Because h is the conjugacy, we know that $h \circ g_1 = g_2 \circ h$. Substituting into the above equation gives

$$\mathfrak{C}_{g_1}^{g_2} h(x) = \hat{g}_{2i}^{-1} \circ g_2 \circ h(x) = (\hat{g}_{2i}^{-1} \circ g_2)(y). \tag{43}$$

Because $y \in I_{Yi}$, and $(\hat{g}_{2i}^{-1} \circ g_2)|I_{Yi}$ is equivalent to the identity, we have

$$\mathfrak{C}_{g_1}^{g_2} h(x) = y = h(x).$$

□

Although h is a fixed point of the operator, we have not established that we can use an iterative scheme to find h. In the best situation, we would like the commutation operator to be a contraction (as it was for the class of tent maps considered in §4 4.1). In application to maps with chaotic attractors, we have found that iteration under the operator numerically converges to a fixed point function.

6.2. The *commutation operator* for maps that are not (necessarily) conjugate.

We now consider the problem of creating a commutation operator when map g_1 and g_2 are not necessarily conjugate. We would like this operator to be a contraction, so that its fixed point could be found by fixed point iteration. Moreover, we want the fixed point to satisfy the commutative diagram. In this section, we will simply address the problem of creating the operator such that the fixed point of the operator satisfies the commutation diagram on some positive measure subset. The primary difference between this problem and that of the previous section is that because the systems are not conjugate, there is no *a priori* reason that we should be able to match monotonic intervals of one dynamical system to monotonic intervals of the other. Consequently, the method requires some *ad hoc* component that specifies this matching. As in the previous section, the underlying principle is that we will construct an operator whose fixed point satisfies the commutative diagram. If the resultant operator is a contraction, then the fixed point can be found by an iterative method. In general, g_1 and g_2 are not conjugate, which means that the fixed point is not a conjugacy. We will use *commuter* to denote a fixed point of this general commutation operator, where the term appropriately describes that the fixed point satisfies the commutative diagram.

To motivate our methodology, we start with the "idea" that the commuter, $f : X \to Y$, should act like a change of variables, associating a trajectory of dynamical system g_1 with a trajectory of g_2. In mapping an *interval* $I_x \subset X$ to a $f[I_x] \subset Y$, we would really like to be able to associate trajectories of points in I_x to trajectories of points in $f[I_x]$. If the systems are not conjugate, then the "matching" will not be perfect.

When the two systems *are* conjugate, the commuter is continuous, mapping intervals to intervals, and we are able to easily indentify an appropriate matching by the maximum intervals on which the dynamical systems are monotone. In essence, it is this matching which prescribes the operator.

Construction of the commutation operator requires that the mathematician, now acting as a modeler, make some choices regarding the particular matching scheme to be employed. To describe our technique, we will decompose the problem into two components: first, we define to be a set of *minimal requirements* to constructing the operator; secondly, we define a set of *recommended choices* for the construction, where we provide some justification for those recommendations.

The following are requirements for construction of a commutation operator:

1. Choose an integer n, which will be the number of intervals for which we will prescribe a matching.
2. Choose a collection of disjoint subintervals $\{I_{Xi}\}_{i=1}^{n}$, where $I_{Xi} \subset X$, and $I = \cup_i I_{Xi}$ is a closed set which is forward invariant under g_1. Without loss of generality, we assume that these intervals are ordered, such that if $x_1 \in I_{Xi}$ and $x_2 \in I_{Xj}$, then $x_1 < x_2$ whenever $i < j$.
3. For each $i \in \{1, \ldots, n\}$, assign inverse function $\hat{g}_{2i}^{-1} : Y \to Y$ which satisfies:
   - $\hat{g}_{2i}^{-1}$ is continuous on Y,
   - There is an associated interval $I_{Yi} \subset Y$, such that $\hat{g}_{2i}^{-1}$ acts like an inverse to g_2 on the interval I_{Yi}. Equivalently, for all $y \in I_{Yi}$, $\quad \hat{g}_{2i}^{-1}(g_2(y)) = y$.
   - $\hat{g}_{2i}^{-1}$ is Lipshitz continuous on $Y - g_2[I_{Yi}]$, with Lipshitz constant $L < 1$.

Using the above minimal choices, the definition of *commutation operator* given in (41) remains notationally correct. However, we need to recognize that the construction depended upon the actual choices for $\{I_{Xi}\}$ and $\{\hat{g}_{2i}^{-1}\}$, and we should formally treat them as parameters in describing how the operator acts on functions:

$$\mathfrak{C}_{g_1}^{g_2} F(x) \equiv \mathfrak{C}_{g_1}^{g_2}(\{I_{Xi}\}, \{\hat{g}_{2i}^{-1}\})F(x) := \hat{g}_{2i}^{-1} \circ F \circ g_1(x) \qquad x \in I_{Xi}. \tag{44}$$

Based on our experience, we provide the following list of *recommendations* for formulating the commutation operator, where our goal is to improve the utility of the resultant fixed point. In particular, if the systems are conjugate, we would like the fixed point to be the required conjugacy. When the systems are not conjugate, we still would like that it act as a commuter from X to Y, while retaining as much of the character of a homeomorphism as possible. We recognize that in some modeling situations there will be good reasons to take alternative approaches to those outlined below.

- *Choose n to be as small as practical.* As n increases, the modeler is required to make more choices regarding which intervals of X should be matched with which intervals of Y. Unless the modeler has an *a priori* reason to force a particular matching, it is better to allow the algorithm to find a matching. Consequently, it is generally counter-productive to choose n to be larger than n_2, where n_2 is the number of monotone sequences of g_2. On the other hand, g_2 must be monotone on each interval I_{Yi}, to allow the inverse to be defined. So if the graphs of $\hat{g}_{2i}^{-1}$ are to cover [29] the graph of g_2, we will need at least n_2 different inverse functions, requiring $n \geq n_2$. Therefore, we typically choose $n = n_2$.
- We usually require $I \equiv X$. As a slight weakening of that condition, we might simply require that I be a closed interval in X. The basic idea is that we would like the commuter to map X to Y, so we need to cover as much of X as is possible.
- If $n \geq n_2$, we will construct $\hat{g}_{2i}^{-1}$ so thatI_{Yi} are disjoint and that $\cup_i I_{Yi}$ is a closed interval. Typically, $\cup_i I_{Yi} \equiv Y$. This construction allows the covering of g_2 described above. Additionaly, because the I_{Xi} have been ordered, we assume that $\hat{g}_{2i}^{-1}$ have been chosen to reflect a similar ordering on the I_{Yi}.

Remark: For the divergence measurements of the commutator f discussed in the next section, we would like to have at least existence and uniqueness of f in all of the general settings just discussed. Our experience indicates that a useful f does exist uniquely for all of the widely varied modeling choices of domains, partitions, and nonconjugate systems we have made. For the time being however, proof of the generality of our method remains open, although our remarks are suggestive of the methods of proof designed for specific examples may generalize. Furthermore, it remains open to prove that we can always construct a general f. Notice that in the next section, we begin under the assumption that f exists, we have it, and we plan to measure it.

7. MEASURE OF *MOSTLY CONJUGATE*

Suppose we have two dynamical systems,

$$\begin{aligned} g_1 &: X \to X, \\ g_2 &: Y \to Y. \end{aligned} \tag{45}$$

When the two systems are topologically conjugate, then the dynamics of one system completely describe the dynamics of the other. However, if they are not conjugate, we may find that "some" of the dynamics of one system can be described by the other. We might heuristically judge that "most" of the dynamics are well represented by the other system. In some sense, we could consider the two systems to be "close," but that description is only reasonable if we can describe a way of "measuring" the "distance from conjugacy." As an application of this "distance from conjugacy measurement," we consider this prototypical problem: Fix g_1, and consider some family $\mathcal{D}$ of dynamical systems; find the element $g^* \in \mathcal{D}$ that most closely approximates the *dynamics* of g_1. Topological conjugacy defines two systems as having the same dynamics, and any notion of "distance to conjugacy" ought to provide a means of determining the extent to which the dynamics are similar. *In this section, we propose a construct that allows such a measurement.*

In Sec. 6, we described a technique for finding a *commuter*, f, from g_1 to g_2. Although f satisfies the commutative diagram, it need not be a homeomorphism (and, therefore, not a conjugacy). Let $\lambda(f)$ be a measure of how far f deviates from a homeomorphism, with $\lambda(f) \equiv 0$ when g_1 and g_2 are conjugate. In Sec. 8, we give a more thorough description of λ, but for now, we simply state that the measuring function λ should be defined with sufficient flexibilty to measure f relative to the "important" subset of X and Y, where "important" and the associated measure would be problem dependent. If we define $\mathcal{CF}(g_1, g_2)$ to be the set of all commuters from g_1 to g_2, then for a given λ, we can define

$$\delta(g_1, g_2) \equiv \inf_{f \in \mathcal{CF}(g_1, g_2)} \lambda(f). \tag{46}$$

As an expository description of δ, we may think of a commuting function f as taking us from dynamical system g_1 to g_2, and $\lambda(f)$ measures a cost. Then δ would describe a greatest lower bound for that cost. In general, we would not require that the cost function be symmetric [30] , so that $\delta(g_1, g_2)$ need not be the same as $\delta(g_2, g_1)$. However, we will require that $\delta = 0$ if g_1 and g_2 are conjugate.

Although δ appears to be a useful *theoretical* construct, we have no method for performing the optimization to evaluate. Therefore, we also define a less global definition by taking the following approach: In Sec. 6, we described techniques to construct the *commutation operator*, $\mathfrak{C}$ for a given g_1 and g_2. Suppose that within a certain class of problems, we choose a particular cannonical technique to construct the commutation operator such that given a particular choice of g_1 and g_2, the operator is *uniquely* defined. This unique choice of operator will have an associated fixed point, $f_{\mathfrak{C}}$ which is a commuter. Then we can use this commuter to evaluate the cost of transforming between the dynamical systems, defining

$$d(g_1, g_2) \equiv \lambda(f_{\mathfrak{C}}). \tag{47}$$

In practical application, given a dynamical system g, we may wish to restrict the comparison to dynamical systems $\tilde{g}$ from a particular family $\mathcal{G}$. Some examples of such restricted families include:

- skew tents
- symmetric tents
- constant slope maps
- polynomial maps of degree n.

The point of using these restricted families is that (1) it may be easier to provide a methodology for producing *unique* commutation operators; and (2) we may choose families that are described by a finite set of free parameters, making it easier to search within that family (3) the properties of the canonical system may be universally understood. When applying this restriction to the general problem of comparing an arbitrary g_1 and g_2, it is then necessary to project the problem onto the family $\mathcal{G}$. We denote $\bar{g}$, the projection of g into $\mathcal{G}$ by

$$\overline{g} = \arg\min_{\tilde{g} \in \mathcal{G}} d(g, \tilde{g}). \tag{48}$$

From a theoretical aspect, we need to be concerned with both the existence and uniqueness of this minimizer. In this paper, we will simply argue that our approach is meant to give a computeable approximation, and within that framework of numerical implementation, it appears that our projection operation is sufficiently robust. Using this projection, we now define a new measurement function $D_{\mathcal{G}}$ by

$$D_{\mathcal{G}}(g_1, g_2) = d(g_1, \overline{g_1}) + d(\overline{g_1}, \overline{g_2}) + d(g_2, \overline{g_2}). \tag{49}$$

We make the following remarks regarding $D_{\mathcal{G}}$:

- If $g_1 \in \mathcal{G}$ then we may choose $\overline{g_1} = g_1 \Rightarrow d(g_1, \overline{g_1}) = 0$.
- If $g_2 \in \mathcal{G}$ then we may choose $\overline{g_2} = g_2 \Rightarrow d(g_2, \overline{g_2}) = 0$.
- If $g_1, g_2 \in \mathcal{G}$, then $D_{\mathcal{G}}(g_1, g_2) \equiv d(g_1, g_2)$.
- If $\mathcal{G}$ is a sufficiently large family such that g_1 and g_2 can be (arbitrarily) well approximate, then we may expect $d(g_1, \overline{g_1}) = O(\epsilon), d(g_2, \overline{g_2}) = O(\epsilon)$, and

$$D_{\mathcal{G}}(g_1, g_2) = d(\overline{g_1}, \overline{g_1}) + O(\epsilon).$$

We also recognize that an alternative approach in choosing projections $\overline{g_1}$ and $\overline{g_2}$ would be to choose a projection pair that minimizes the sum, $d(g_1, \overline{g_1}) + d(\overline{g_1}, \overline{g_2}) + d(g_2, \overline{g_2})$. However, that approach might add significant computational complexity to the optimization problem and we have found the simpler approach to be sufficient in application.

7.1. Manifesto for our particular choice for $\mathcal{G}$.

Consider the family of dynamical systems $\mathcal{T}$ defined by:

- Each $f \in \mathcal{T}$ is continuous and piecewise linear on $[a, b]$.
- For each $f \in \mathcal{T}$, $|T'(x)| = c$, constant wherever the derivative exists.
- For each $f \in \mathcal{T}$, $|T'(x)|$ exists for all but perhaps a finite set of points.

We call $\mathcal{T}$ the *the set of constant slope maps.* In several applications, we note that choosing $\mathcal{G} \subset \mathcal{T}$ results in some nice properties for the projections because they inherit these properties from dynamical systems characteristics of constant slope maps. We briefly (conjecture) outline below:

- Metric entropy, h_m, over any invariant measure, satisfies

$$h_m = \ln |f'| = h_T,$$

  where h_T is topological entropy.
- If $|f'(x)| = c > 1$, the map is everywhere expanding.
- For all maps with $|f'| > 1$, the dynamical system has an ergodic measure whose support has positive Lebesgue measure.

8. MEASURING THE DEVIATION FROM HOMEOMORPHISM

Suppose we have two dynamical systems, $g_1 : X \to X$ and $g_2 : Y \to Y$. Additionally, suppose $f : D \to R$ is a commuter, such that

$$f \circ g_1 = g_2 \circ f, \tag{50}$$

as represented in the commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{g_1} & X \\ {\scriptstyle f}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ Y & \xrightarrow{g_2} & Y. \end{array} \tag{51}$$

If f were a homeomorphism, then g_1 and g_2 would be conjugate. However, if we know that the two dynamical systems are not conjugate, then f must fail to be a homeomorphism. We desire to build a metric which measures the extent to which f fails to be a conjugacy. We note that f would fail to be homeomorphic if

- It is not onto.
- It is not one-to-one.
- It is not continuous.
- Its inverse is not continuous.

Our general strategy will be to define a *homeomorphic defect*, which provides a weighted average based on measurements of each of these possible failures. We denote

$$\begin{aligned} \lambda_O(f) &= \{ \text{ amount that } f \text{ is not onto } \}, \\ \lambda_{1-1}(f) &= \{ \text{ amount that } f \text{ is not } 1-1 \}, \\ \lambda_C(f) &= \{ \text{ amount that } f \text{ is not continuous } \}, \\ \lambda_{C^{-1}}(f) &= \left\{ \text{ amount that } f^{-1} \text{ is not continuous} \right\}, \end{aligned}$$

where we note that f^{-1} may not be well defined. While there are certainly many ways to define each of these components, a specific definition scheme should satisfy

- $\lambda_O(f) \geq 0$, with equality when f is onto; [31]
- $\lambda_{1-1}(f) \geq 0$, with equality when f is $1-1$;
- $\lambda_C(f) \geq 0$, with equality when f is continuous; and
- $\lambda_{C^{-1}}(f) \geq 0$, with equality when f^{-1} is continuous.

We define *homeomorphic defect of* f, denoted $\lambda(f)$, as a convex combination

$$\lambda(f) = \alpha_1\lambda_O(f) + \alpha_2\lambda_{1-1}(f) + \alpha_3\lambda_C(f) + \alpha_4\lambda_{C^{-1}}(f), \tag{52}$$

where the weights satisfy

$$0 \leq \alpha_i \leq 1, \text{ and } \sum \alpha_i = 1. \tag{53}$$

Our decision not to specify a particular choice for the weights allows the flexibility to "tune" this metric for a particular application. Then

$$\lambda(f) \geq 0, \text{ with equality when } f \text{ is a homeomorphism.} \tag{54}$$

We note that if the converse were to hold, such that

$$\lambda(f) = 0 \Longrightarrow f \text{ a is homeomorphism,} \tag{55}$$

then λ could be used to provide a "distance from conjugacy" for the two dynamical systems g_1 and g_2. In this paper, our primary goal is to maintain the flexibility of the definitions (to allow broader applicability). Indeed, to retain this flexibility, our definitions are measure based and, consequently, we expect that the converse will not hold.

8.1. Supporting assumptions and notation

As preliminary material, we assume that we have identified measure spaces $(D_1, \mathcal{A}_1, \mu_1)$ and $(D_2, \mathcal{A}_2, \mu_2)$, where $D_1 \subset X$ and $D_2 \subset Y$, $\mathcal{A}_1$ and $\mathcal{A}_2$ are σ-algebras, and μ_1 and μ_2 are measures. D_1 and D_2 are "chosen" by the modeler, and represent the subsets of X and Y that are of interest to the modeler. For example, one might be interested in comparing the dynamics of g_1 and g_2 on their forward invarinant sets, which might be smaller than the whole space of the dynamics. Additionally, by allowing the modeler to specify a measure on the "important" parts of the sets D_1 and D_2 can be more heavily weighted. In some cases, the dynamics of interest might lie on $D_1 = C$, with C an invariant Cantor set with Lebesgue measure-0. We may still choose μ_1 such that $\mu_1(D_1) > 0$, which would allow us to measure subsets of D_1 in ways that will distinguish there size. Typically, we will be interested in chaotic dynamics, and assume that X and Y are bounded sets. In all the examples in this paper, the sets D_1 and D_2 are closed intervals and μ_1 and μ_2 are Lebesgue measures; for simplicity, we call this the *standard case*. Unless otherwise stated, we assume that μ_1 and μ_2 are *finite* and *non-atomic* measures. Additionally, we assume that $D_1 \subset \mathbb{R}^n$ and $D_2 \subset \mathbb{R}^n$, with topologies inherited from the usual topology. Our notation is that $f : D \to R$, is a commutator. Generally, we think of D_1 as being a restriction to a set of interest, but we do not preclude the case that $D_1 \not\subset D$ (or similarly $D_2 \not\subset R$). These situations reflect that sometimes the modeler is interested in a set that is larger than where the commutator is defined. As necessary, we will simply assume that the subsets resulting from various intersections are μ_1 or μ_2 measureable whenever this measureability is required by some definition. For ease of notation, we define

$$\overline{\mu_2}(f[A]) = \mu_2(f[A \cap D_1] \cap D_2),$$

for arbitrary set $A \subset X$. The idea is that we want to restrict ourselves to measuring image points that lie in D_2 whose pre-image was in D_1. Similarly, we define

$$\overline{\mu_1}(f^{-1}[B]) = \mu_1(f^{-1}[B \cap D_2] \cap D_1).$$

8.2. "Onto" deficiency

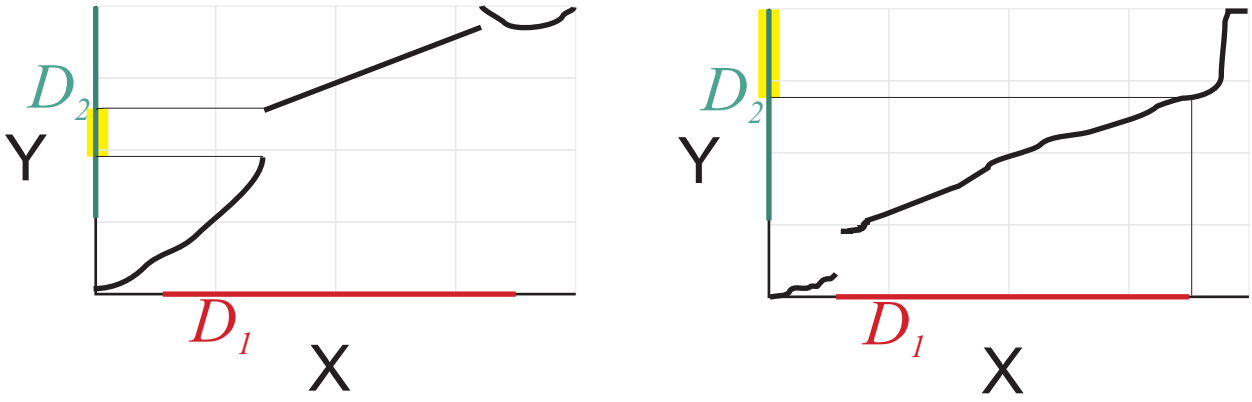


FIG. 6: **Illustration of onto deficiency.** See definition of λ_O, (56). The space X is represented on the horizontal, with Y the vertical. The sets of interest, D_1 and D_2, are colored red and green. On each graph, we show an example of the graph of a commuter, $f(x)$. The "onto-deficiency" is computed from a μ_2 measure of the yellow portion. The righthand example illustrates that (1) we are only interested in measuring the extent to which D_2 is covered (so the lower gap is not a problem), and (2) we are only allowed to use the part of f that is over D_1.

To measure the onto deficiency, we desire to measure the fraction of D_2 which is not-covered by the range of f. We define the *onto deficiency*, λ_O of the function f by

$$\lambda_O(f) = 1 - \frac{\overline{\mu_2}\,(f[D_1])}{\mu_2(D_2)}. \tag{56}$$

See Fig. 6.

Because f may have fractal structure with the range of f a Cantor set, it may be difficult to implement (56) in computational practice. As a "suitable surrogate," we find that if D_2 is an interval and μ_2 is absolutely continuous, we can often quantify the lack of onto-ness by finding the "biggest hole" in the range of f. Specifically, if we define

$$G := D_2 - f[D_1], \tag{57}$$

then a suitable surrogate is given by

$$\tilde{\lambda}_O(f) := \sup_{I \subset G} m(I), \tag{58}$$

where I is an interval and m is Lebesgue measure. Note that generally $\tilde{\lambda}_O(f) \not\approx \lambda_O(f)$, but we expect that the values $\tilde{\lambda}_O(f)$ and $\lambda_O(f)$ will both change in the same direction (either increase or decrease) in response to a change in the argument function f. In other words, the hallmarks of an appropriate *suitable surrogate* is such monotonicity with respect to parameter variations between $\tilde{\lambda}$ and λ. Such may not always hold true, but is often true in application on the simple examples we describe here.

8.3. $1-1$ deficiency

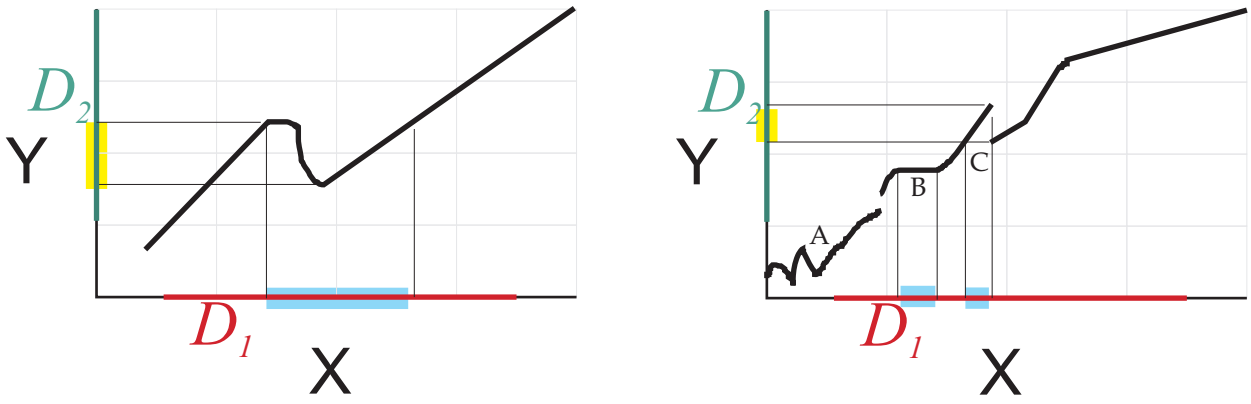


FIG. 7: **1-1 defect.** See definition of λ_{1-1}, (59). (Left) A typical deficiency, where we take a μ_1 measure of the set that must be removed from D_1 to make the remaining function 1-1 (the blue horizontal segment) and a μ_2 measure of the portion of D_2 which is multiply covered (the yellow). (Right) In this example, note: (1) the non-monotonicity near region A is not measured, becuase it lies outside the sets of interest; (2) The horizontal component in region B results in a defect from the horizontal measurement, but no contribution from the vertical; (3) In region C, we take the interval to the left of the jump to perform the horizontal measurement because of the inf in the definition of λ_{1-1}.

To measure the extent to which f is not $1-1$, we need to quantify *where* the function is not $1-1$, by measurement on the domain of f, and the extent of the folding [32] , a measurement on the range. We procede as follows: we define $\mathcal{G}$ to be the collection of all subsets $G \subset D_1$ which satisfy

- G is μ_1 measureable;
- $f[G]$ is $\overline{\mu_2}$ measureable;
- f restricted to G is $1-1$;

For any such G, we denote its complement in D_1 by $\bar{G} \equiv D_1 - G$. Then we define the $1-1$ *defect* by

$$\lambda_{1-1}(f) := \inf_{G \subset \mathcal{G}} \left[\frac{\mu_1(\bar{G})}{2\mu_1(D_1)} + \frac{\overline{\mu_2}(f[\bar{G}])}{2\mu_2(D_2)} \right]. \tag{59}$$

See Fig. 7.

In the *standard case,* we simply try to identify the largest part of the range that is multiply covered. We define envelope functions

$$e^+(x) = \sup_{D_1 \ni y \le x} f(y) \tag{60}$$

$$e^-(x) = \inf_{D_1 \ni y \ge x} f(y). \tag{61}$$

Then $e^+(x)$ records the largest function value to the left of x, while $e^-(x)$ records the smallest function value to the right of x. Then

$$\tilde{\lambda}_{1-1}(f) = \|e^+(x) - e^-(x)\|_p. \tag{62}$$

We often choose $p = \infty$, yielding the sup norm, but other choices for p may be useful in some situations.

8.4. Measuring discontinuities of f.

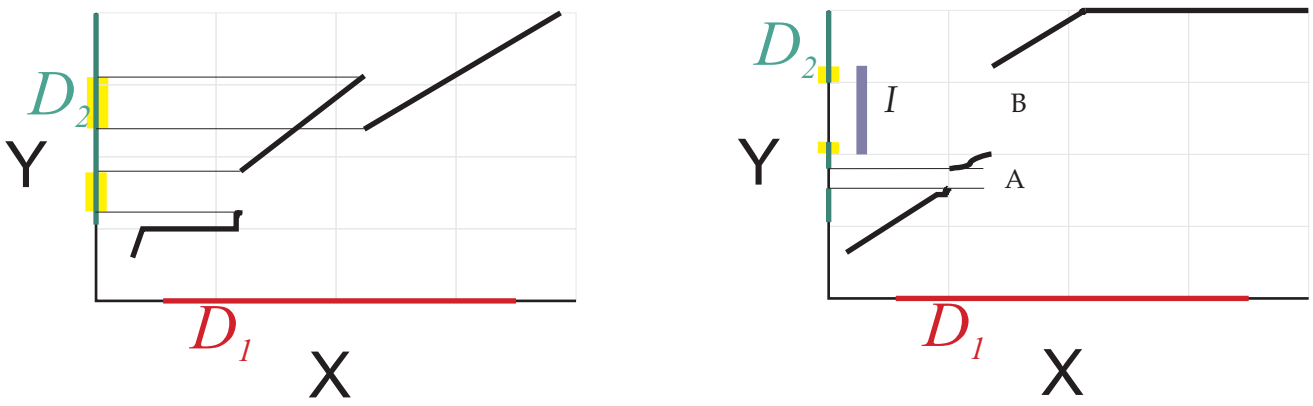


FIG. 8: **Continuity defect.** See definition of λ_C, (66). (Left) For a typical discontinuity, we μ_2 measure the size of the jump (and then choose the largest such jump). (Right) When D_2 is not an interval, we must be more careful. The discontinuity in region A does not create a defect, because it lies outside of D_2, while in region B, we only measure the portion of I that lies inside D_2.

To measure discontinuities, we note that a continuous function maps "small sets to small sets," and that a discontinuity is indicated when an arbitrarily small set is mapped to a large one. In particular, we seek to identify when the f undergoes a "jump," and measure the jump. To allow the possiblity of Cantor sets for D_1 and D_2, we will need to define these ideas in terms of sets and measures of those sets.

For each $x_0 \in D_1$ and for each $\delta > 0$, we define the set

$$\mathcal{B}(\delta, x_0) := \{x : x \in D_1, |x - x_0| < \delta\}, \tag{63}$$

which creates a nested family of sets as $\delta \searrow 0$. We measure the f-image of these sets by defining

$$a_\delta(x_0) := \inf_{I \supset f[\mathcal{B}(\delta, x_0)]} \frac{\mu_2(I \cap D_2)}{\mu_2(D_2)}. \tag{64}$$

Because $a_\delta(x_0)$ is monotonically decreasing with decreasing δ, we can take the limit as $\delta \searrow 0$, defining

$$a(x_0) := \lim_{\delta \to 0^+} a_\delta(x_0), \tag{65}$$

where we think of $a(x_0)$ as being the *atomic* part of f [33] .We define

$$\lambda_C(f) := \sup_{x_0 \in D_1} a(x_0). \tag{66}$$

See Fig. 8.

Because $\lambda_C(f)$ is fundamentally based on intervals, we generally find that it is sufficiently easy to approximate such that we have not used a surrogate. However, we note that if D_1 is an interval, we can define

$$\tilde{\lambda}_C(f) = \|a(x_0)\|_p. \tag{67}$$

We note that $\tilde{\lambda}_C(f) \equiv \lambda_C(f)$ when $p = \infty$, but the flexibility to use other norms might prove useful in some situations.

8.5. Discontinuity in f^{-1}.

In the *standard case*, we note that if f is $1-1$ and continuous, then f^{-1} is well defined and is also continuous, so this requirement on a conjugacy may seem redundant. Generally, this requirement for a topological conjugacy is needed because the domain space and range space may be defined with very different topologies. In our situation, although the topologies are typically similar (based on inheriting the usual topology of the line), we find that directly measuring this fourth defect is the easiest and most direct way to measure gaps in the domain of definition. To define this defect, we employ the same strategy as for measuring the continuity of f.

For each $y_0 \in D_2$ and for each $\epsilon > 0$, we define the set

$$\mathcal{B}(\epsilon, y_0) := \{y : y \in D_2, |y - y_0| < \epsilon\}. \tag{68}$$

We measure the pre-image of these sets by defining

$$\hat{a}_\epsilon(y_0) := \inf_{I \supset f^{-1}[\mathcal{B}(\epsilon, y_0)]} \frac{\mu_1(I \cap D_1)}{\mu_1(D_1)}. \tag{69}$$

Taking the limit $\epsilon \searrow 0$, we define

$$\hat{a}(y_0) := \lim_{\epsilon \to 0} \hat{a}_\epsilon(y_0), \tag{70}$$

where we think of $\hat{a}(y_0)$ as being the *atomic* part of f^{-1}. We define

$$\lambda_{C^{-1}}(f) := \sup_{y_0 \in D_2} \hat{a}(y_0). \tag{71}$$

In the *standard case*, we use this defect to measure gaps in the domain. Similarly as for λ_C, a suitable surrogate is to measure the largest such gap in the domain. We define $\mathcal{I}$ to be the set of all intervals $I \subset D_1$ such that $f(x)$ is undefined or constant for all $x \in I$. Then we measure the defect as

$$\tilde{\lambda}_{C^{-1}}(f) = \sup_{I \subset \mathcal{I}} m(I), \tag{72}$$

where m is Lebesgue measure.

The picture for $\lambda_{C^{-1}}(f)$ is not shown, but would be similar to that shown for $\lambda_C(f)$ in Fig. 8, where the roles of domain and range are reversed in the obvious manner for inverse functions.

9. EXAMPLES

In this section, we present several examples of comparing maps g_1 and g_2 by showing the resulting commuter f, whether it be a homeomorphism or not. Each example follows the same presentation template: in the left panel we graph g_1 in blue and g_2 in red. Additionally, the graphs for $\hat{g}_{2i}^{-1}$ are displayed in green circles. Along the horizontal axis, the blue and red rectangles indicate the chosen intervals I_{Xi} and I_{Yi}. The right panel graphs the resultant commuter function created by repeated application of the commutation operator. For each commuter, the caption shows the approximate computation $\tilde{\lambda}(f) = \frac{1}{4}(\tilde{\lambda}_O + \tilde{\lambda}_{1-1} + \tilde{\lambda}_C + \tilde{\lambda}_{C^{-1}})$.

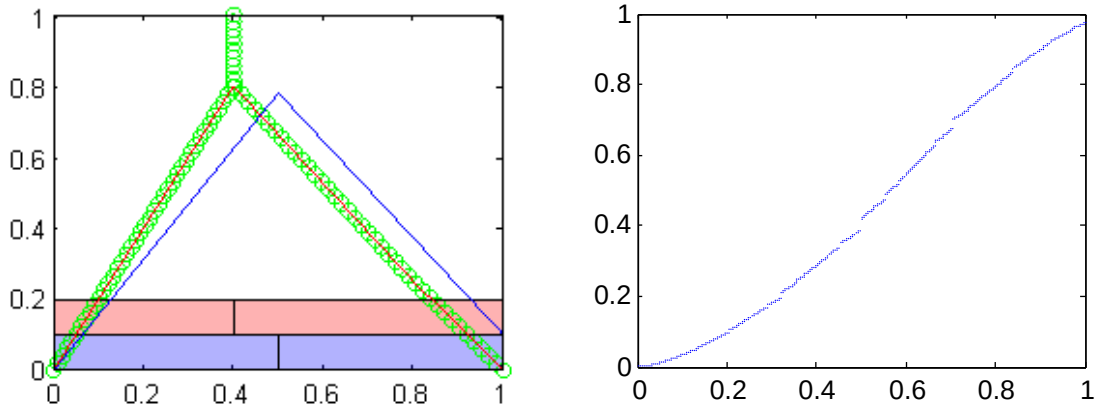


FIG. 9: These two tent maps are almost conjugate. $\tilde{\lambda}(f) \approx \frac{1}{4}(0.032 + 0 + 0.032 + 0)$. Note that $g_1(1) \neq 0$.

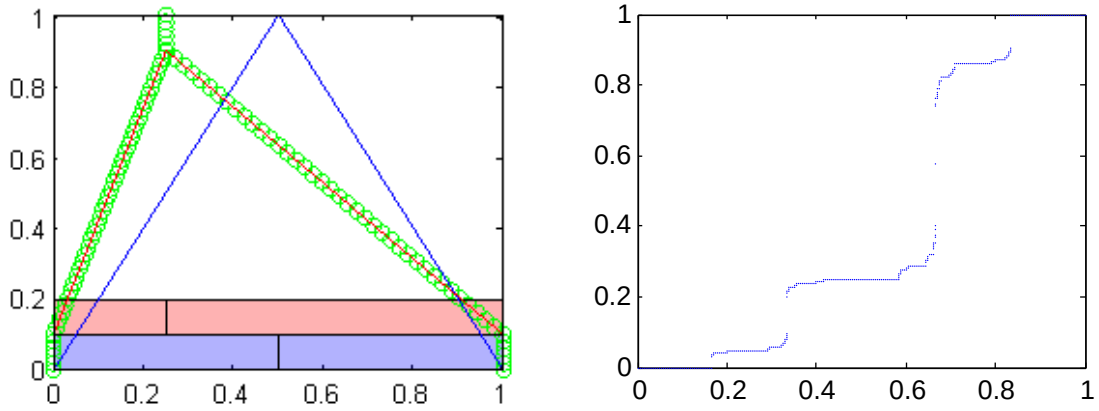


FIG. 10: Maps that are not close to conjugacy. $\tilde{\lambda}(f) \approx \frac{1}{4}(0 + 0 + 0 + 0.163)$. Near $x = 1/2$, the commuter, f is actually horizontal, indicating that there is an interval of orbits of g_1 than cannot be represented in g_2, generating the onto defect. These points are associated with the vertical sections of the green graphs of g_2, which we defined as $\hat{g}_{2i}^{-1}$, which "complete" the inverse of g_2, as required by (41). The apparent vertical gap is not a defect of continuity, but is simply the result of an extremely steep section in the graph, and not enough points are plotted to fill in the picture. At the fixed point of the dynamical systems, the slope of the commuter is infinite, as per (6). We know there is no actual vertical gap because g_1 is a full shift and is able to match all orbits of g_2.

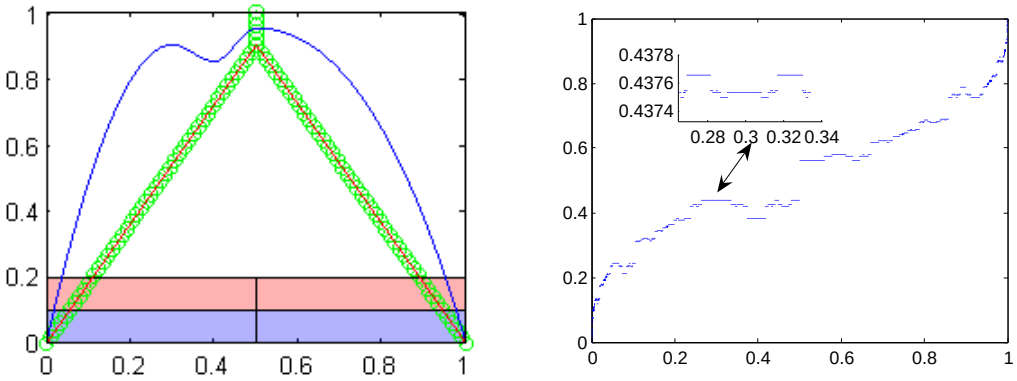


FIG. 11: **Logistic map with "just a little" extra dynamics.** These maps are not as far from conjugacy as appearances might indicate. That g_1 has two humps, and g_2 has one hump, which suggests that subshifts of Σ_4 and Σ_2 respectively are required to represent the dynamics. $\tilde{\lambda}(f) \approx \frac{1}{4}(0.12 + 0.056 + 0.12 + 0)$. In fact, g_1 behaves as if it is almost conjugate to a trapezoid map [20], also known as a gap-map [21], since it does not take a large perturbation to replace the extra humps of $x \in (0.25, 0.5)$, with a horizontal line segment. However, our first choice, using a tent map of approximately the same height as g_1 does not match very well. Note the vertical gaps in the commuter, indicating that g_2 has dynamics that are not matched by g_1.

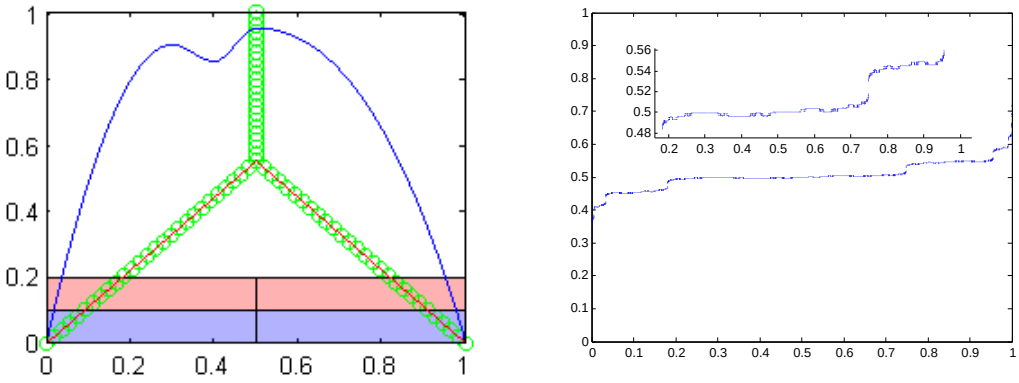


FIG. 12: **Better projective comparison.** The blue map g_1 is identical to that shown in Fig. 11, but here it is compared to a shorter tent map, $g_2(x) = T_{0.55}$. Therefore this g_2 involves much less folding of the interval into itself and is much further from g_1 in an $L^1([0,1])$ sense. Because of the L^1 difference, we might presume g_2 is less close to g_1 in an almost conjugacy sense. Our almost conjugacy surrogates measure $\tilde{\lambda}(f) \approx \frac{1}{4}(\epsilon_1 + 0.004 + \epsilon_3 + 0.031)$, where ϵ_i denotes a numerically small value. Although the graph seems to indicate a large onto deficiency, the small $\tilde{\lambda}_O$ can be understood by first considering only the invariant sets of each map, $D_1 = [0.201, 0.95]$ and $D_2 = [0.495, 0.55]$, instead of the unit intervals $[0,1]$ shown. This scenerio is emphasized by the inset picture of f, where we can see that $\tilde{\lambda}_O$ is small, and the other defects are even smaller. Comparing the maps g_1 and g_2 outside of the invariant sets, say on the left sides, near $x = 0$, the maps are very similar: each sweeps initial conditions monotonically into D_1 and D_2 respectively. Thus the extension of f to the full $[0,1]$ to $[0,1]$ does not suffer the apparent onto deficiency near $x = 0$ and $x = 1$. $\tilde{\lambda}_{1-1}$ is small because this surrogate measures only the difference between the upper and lower envolope, as per (62). f is not continuous, but $\tilde{\lambda}_C$ is small since the vertical steps are short. The largest defect measurement, $\tilde{\lambda}_{C^{-1}}$, comes from short horizontal components, which are measured as discontinuities in f^{-1}. Why would the much shorter tent map g_2 here measure better as almost conjugate than the g_2 shown in Fig. 11? The result is reasonable when we focus on invariant sets: The blue curve on its invariant set is much more like the green g_2 in this example than the g_2 in Fig. 11, where a large fraction of the unit interval is folded over itself. We conclude that although the blue curve is tall, on its invariant set, it appears to be much like a scaled, almost conjugate, version of the short tent map on its invariant set.

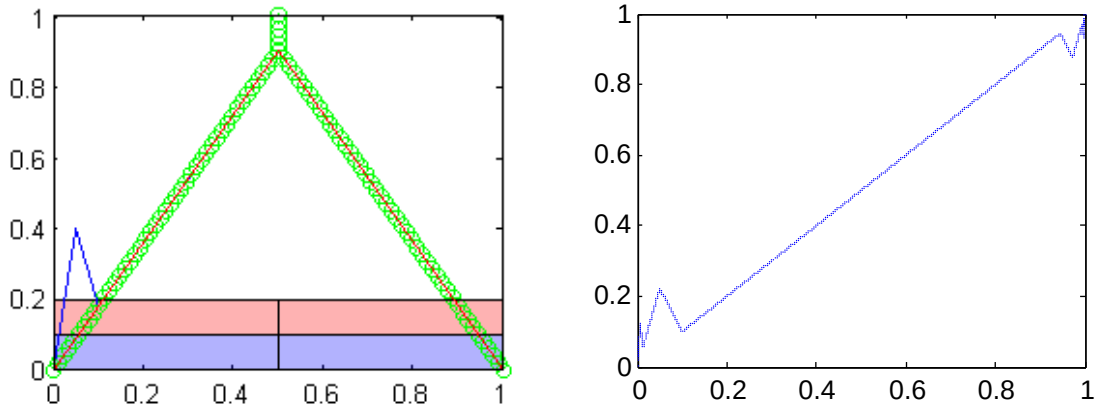


FIG. 13: **Identical on the invariant set.** For this example, g_1 and g_2 are identical on the interval $[0.1, 1]$. However, whereas g_2 is a simple tent map, g_1 has a an additional hump on the interval $[0, 0.1]$. Because the maps are the same over a portion of the domain, the commuter coincides with the identify map on an interval. $\tilde{\lambda}(f) \approx \frac{1}{4}(0+0.12+0+0)$. If dynamics were restricted to the invariant set of each system, then the two maps would be conjugate, with $f(x) = x$ the homeomorphism.

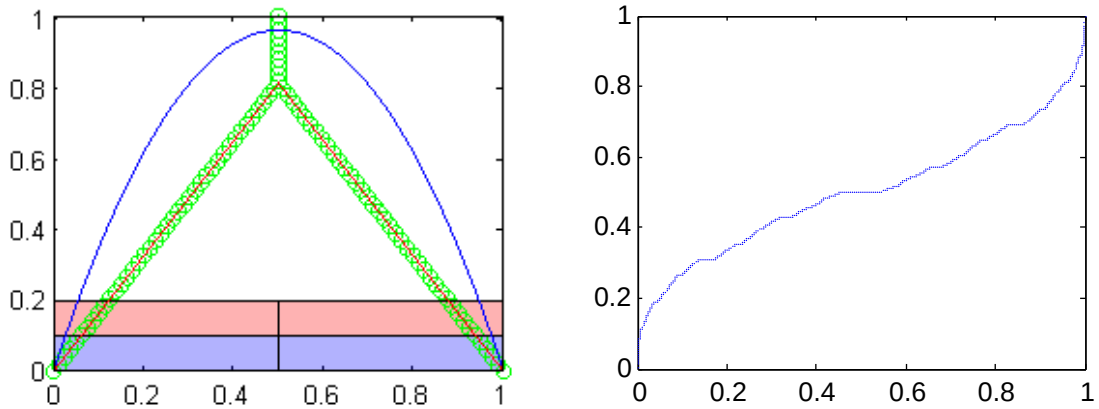


FIG. 14: **The period three window.** Choosing g_1 to be the logistic map with parameter 3.84, there is an attracting period three orbit whose basin is a.e. point in the interval. A Cantor set of initial conditions which are not in that basin constitute a chaotic saddle of g_1. We compare this map to a submaximal tent map, where we have chosen the height of that tent map to so that the defect will be small, with $\tilde{\lambda}(f) \approx \frac{1}{4}(0.0005 + 0 + 0.0005 + 0.022)$. A higher tent map would create large intervals on which f is horizontal, while choosing a smaller tent would create larger vertical gaps. Each horizontal component indicates that f is associating an interval of g_1 dynamics to a single point in g_2.

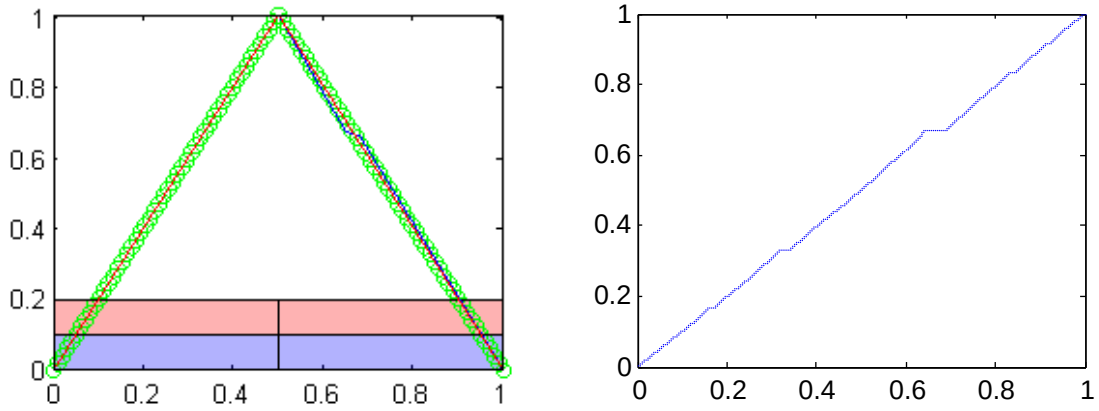


FIG. 15: **Two full-shift maps that are not conjugate** g_1 is a slight alteration of the full-shift symmetric tent. The right leg has been modified so that a small portion has a slope m, with $|m| < 1$, where that interval contains the fixed point. Therefore, g_1 has a stable fixed point at $x_f = 0.665$. This commuter is similar in character to that of the Fig. 14 — the stable periodic point of the X dynamics means that an entire interval of initial conditions must be associated with a single point in the Y dynamics. Since a.e. initial condition is in the basin of attraction for this fixed point, the resultant commuter is a devil's staircase function. $\tilde{\lambda}(f) \approx \frac{1}{4}(0 + 0 + 0 + 0.048)$. We remark that when measuring the defect via suitable surrogate, the deficiency caused by horizontal portions of the commuter are recorded as a defect in the continuity of the inverse, not by the defect in $1 - 1$.

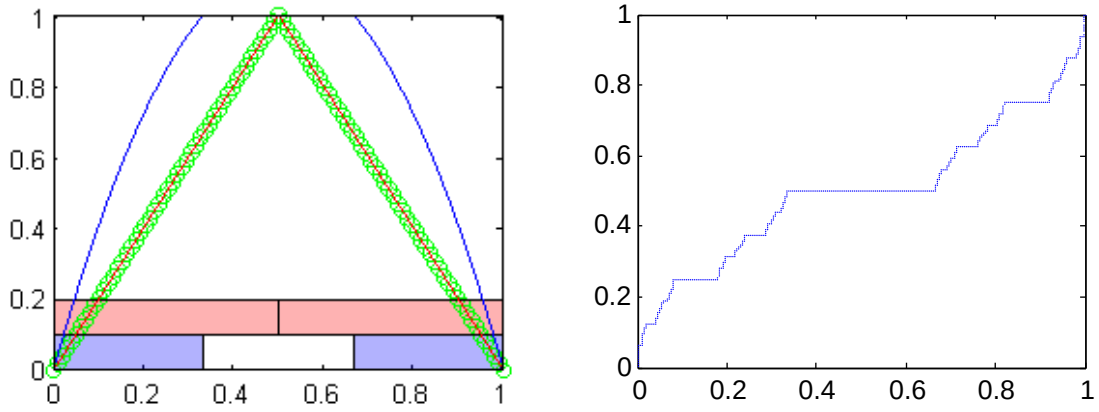


FIG. 16: **The "too-tall" logistic map.** The literature [6, 11] tells us that the invariant set of this logistic map is semi-conjugate to the full tent map. This relationship is a conjugacy if the domain of the tent map excludes the peak point, and all of its preimages. This commuter f, shown here for the first time, has defects $\tilde{\lambda}(f) \approx \frac{1}{4}(0+0+0+1/3)$, is, in fact, a devil staircase function with flat spots over the holes of a Cantor set. Notice the strong similarity to the previous two examples.

10. CONCLUSIONS

In this paper we have put forward a generalizable method based on functional fixed point iteration to explicitly construct the change of coordinates function which acts as a conjugacy between two topologically conjugate dynamical systems. We expect that our method of fixed point iteration will extend, in straight forward manner, to building conjugacy functions between higher dimensional systems when the symbol generating partitions are available. Of course, we know that this last caveat [3, 22] is nontrivial, and this research is one aspect of our continuing work in the subject.

While construction of conjugacy functions is interesting in its own right, we do not consider this to be the main intellectual contribution of the work. In our opinion, a remarkable aspect of our methods of fixed point iteration is that it still produces a commuter, even between nonequivalent systems, and in a sensible manner according to choices and preconceptions of the modeler. We put forward that the major contribution of this work is the thesis that dynamical systems should be comparable in a sense which is consistent with the topological conjugacy notion of equivalence; this is, after all, the centerpiece of the field of dynamical systems. Whereas nonequivalent systems have typically been compared by choice of a particular functional Banach space, such as L^2, it is obvious and well known that similarity or dissimilarity within such a norm does not directly address comparison between orbits of the two systems. Instead, we have claimed that comparison of nonequivalent dynamical systems could be made by via the commuter function by measuring how it fails to be a homeomorphism. We believe that the specific details of our defect cost function λ are reasonable and useful to our stated purposes in the manifesto of Section . Furthermore, we have designed flexibility for the modeler to design the details of λ so as to focus on various aspects of the mismatch. As we have shown in the examples of Section , a nonzero defect is generally descriptive of nonmatching between some of the orbits of the two dynamical systems, and the *concept* of measuring mismatch is the main issue we wish to emphasize, separate from the details of our defect functions. Thus, beyond entropy, we have outlined a new method measuring the dissimilarity of the languages corresponding to orbit structures of two nonequivalent dynamical systems. In this sense, our new methods may be considered more natural, within the context of a dynamical systems based comparison, than a direct comparison of the right hand sides of dynamical systems by using the norm of the difference in some Banach space.

There are a number of theoretical and applications oriented directions that we are currently pursuing in the future development of this work. These efforts include a topological based parameter estimation scheme for modeling dynamical systems within an assumed model classes, for example to match "toy models" to observed data. Similarly, our methods should be considered as a principled way to validate models produced by analysis of time-delay embedded data, where the primary desire is proper representation of the dynamical characteristics of the original system. Furthermore, since certain systems of differential equations (such as the Lorenz equations) admit Poincare' surface of sections which are very much like one-dimensional maps, our techniques are already directly applicable to such ODEs. We are also pursuing extension of our fixed point iteration scheme to allow for multivariate systems, to within our ability to decide generating partitions. In this context, there is promise to extend our methods for comparison of general classes of differential equations. There is no fundamental road block to a multivariate extension of the fixed point iteration scheme, and since the design of the defect functions was entirely measure based, it will extend directly to this setting. Finally, on theoretical grounds, we understand that our commuter maps often serve as factors between two dynamical systems, and there are useful connections to be made between our commuter maps, and the notion of factor maps one uses within the context of symbolic dynamics theory [13]. A future goal for this work is to be able to measure the degree to which a toy model is descriptive of the larger system (such as a Galerkin projection of a PDE).

Acknowledgments

This research was supported under NSF grant DMS-0404778.

11. APPENDIX: QUADWEBBING

"Quadwebbing" is our name for a graphical representation that allows us to visualize the action of the both commutation operator and the commuter in relation to 1-d maps of the interval. The basic structure

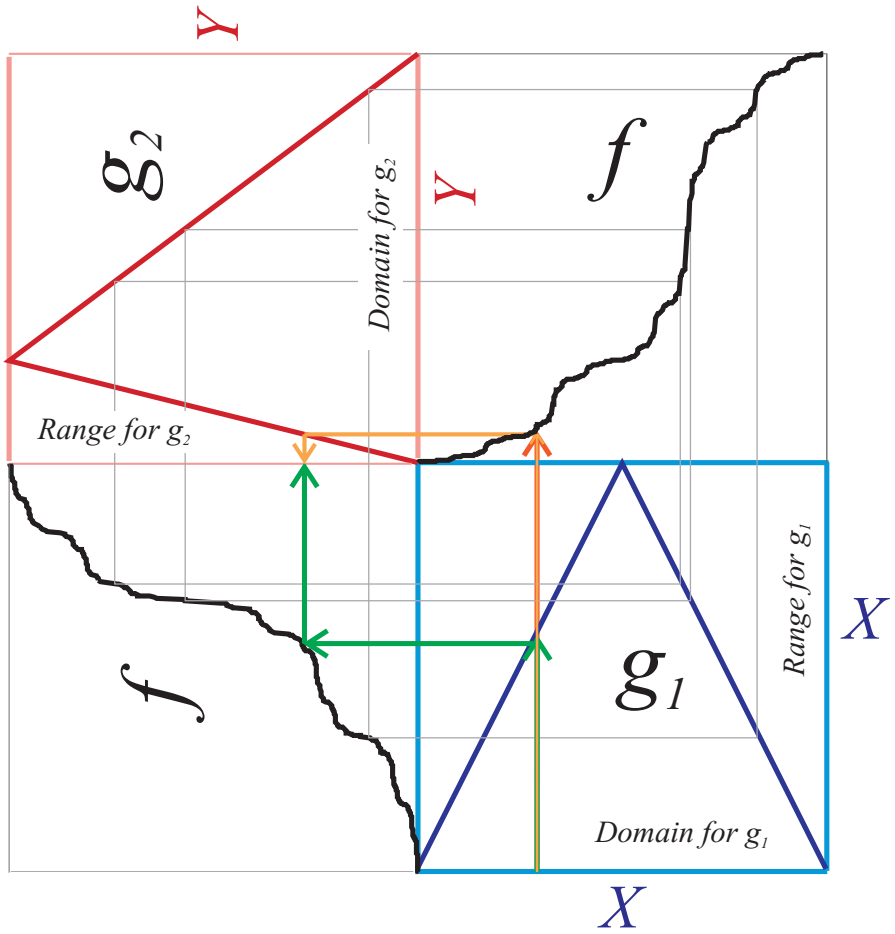


FIG. 17: **Quadweb for conjugate tent maps.** The basic structure for all quadwebs: In the lower right panel, we plot dynamical system g_1 in the normal orientation. Since this maps X to X, all four edges of that panel are colored light blue, to indicate that they all describe X, with the domain on the horizontal, and the range on the vertical. System g_2 is plotted in the upper left, but it is rotated counterclockwise . The edges of that panel (colored pink) all represent Y space, with g_2 mapping the vertical domain to the horizontal range. The commuter f mapping X to Y appears in upper right panel with normal orientation, mapping domain X to domain Y, equivalent to the left side of the commutative diagram. f also appears in the lower left panel, rotated counterclockwise so that it maps range space X to range space Y, the right side of the commutative diagram. The green arrow illustrate the computation of $f \circ g_1(x)$ for a particular x. The orange arrow uses that same x, but illutrates the computation $g_2 \circ f$. Because f satisfies the commutative diagram, the ends of these arrows must meet — $f \circ g_1(x) = g_2 \circ f$. The gray rectangles indicate that this relationship holds for arbitrary x.

of a quadweb diagram is based on the idea that we seek to understand the relationship between dynamical systems g_1 and g_2 within the context of the commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{g_1} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow{g_2} & Y \end{array} \tag{73}$$

Each quadweb is constructed by dividing the figure into a 2-by-2 panel of axes, where each panel shows the graph of one of the four function shown in the commutative diagram. The plots are arranged to allow a graphical illustration of how the function f satisfies the commutative diagram. Figure 17 illustrates the case of conjugate dynamical systems.

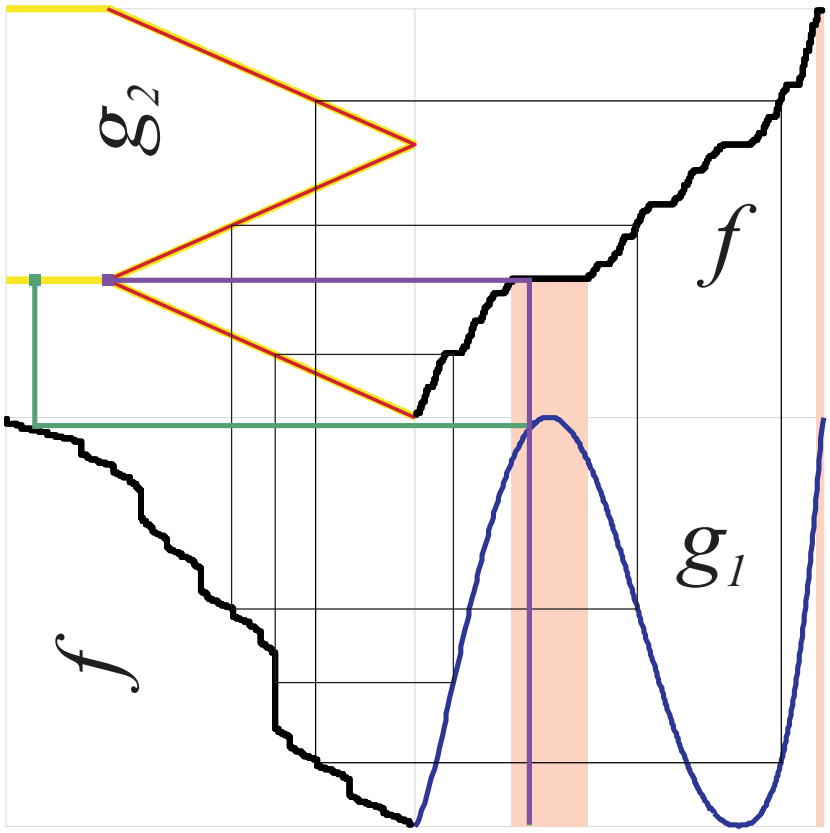


FIG. 18: **Quadweb for non-conjugate maps.** g_1 maps across X a full three times. g_2, shown in red, also has three laps, but they do not map completely across Y. The yellow graph shows the $\hat{g}_{2i}^{-1}$, which are used to construct a well defined commutation operator. The green path illustrates $f \circ g_1$ for a particular x. Using the same x, we use the purple path to show $g_2 \circ f(x)$, which obviously does not coincide with the green. Rather the green path shows that we land on the *zed* which is the case for all x inside the invervals colored pink. The black rectangles illustrate that (73) holds at x lying outside these intervals.

In addition to illustrating how the commuter satisfies (73), it reveals where it might fail. Specifically, we recall from Sec. 6, we noted that to have a well defined operator, we needed to extend the graph of g_2 such that its inverse could be applied to any $y \in Y$. We denoted the modified graphs as $\hat{g}_{2i}^{-1}$. We use the term [26],

The *zed* describes that portion of the graph of $\hat{g}_{2i}^{-1}$ that does not coincide with g_2. (74)

If $f \circ g_1(x)$ lies on the zed, then (73) does not hold for that x. This phenomena can be interpreted to mean that g_1 has more dynamics than can be represented by g_2 under the assumed partitions. Figure 18 illustrate this behavior.

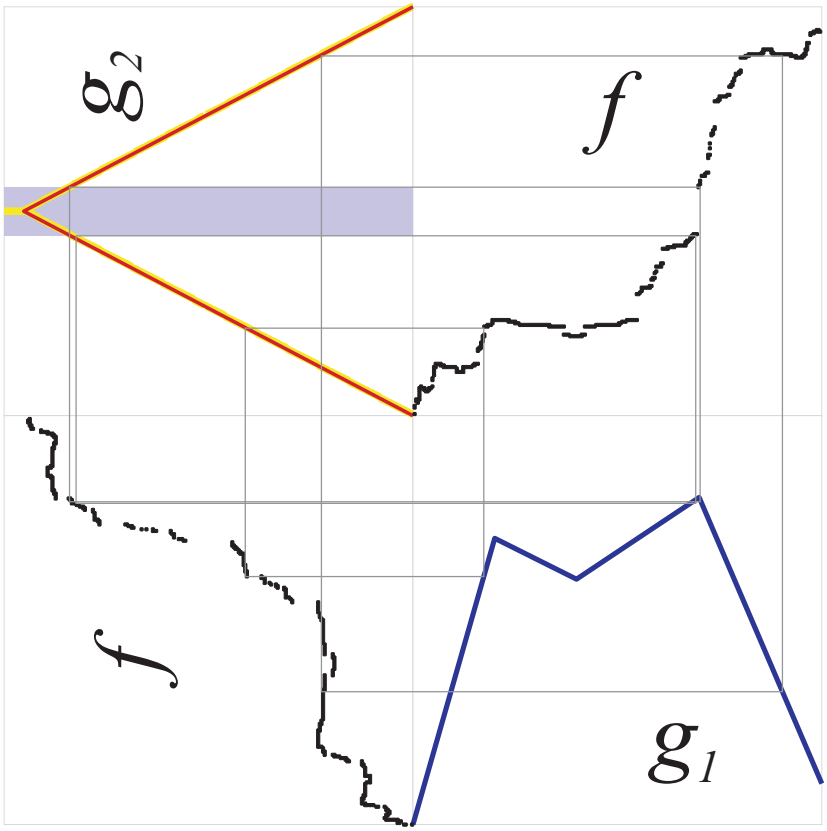


FIG. 19: **Quadweb — Vertical gap in the commuter.** The commuter shown was computed by partitioning g_1 at its maximum. At the x coordinate which is the preimage of that peak, we see a vertical gap in f. Choosing nearby points to the left and right, we will match to g_2 points on either the rising or falling side of the tent, but we cannot match the peak of g_2. The purple rectangle illustrates one of the intervals in g_2 that has no matching dynamics in system g_1.

The quadweb diagram allows for easy understanding of what some typical defects imply about the dynamics of the system. In particular, wherever f is horizontal, we see an interval of initial conditions for system g_1 that must be represented by a single point under g_2. Similarly (see Fig 19) when f has a vertical gap, there is an interval in g_2 that has no associated orbit in system g_1.

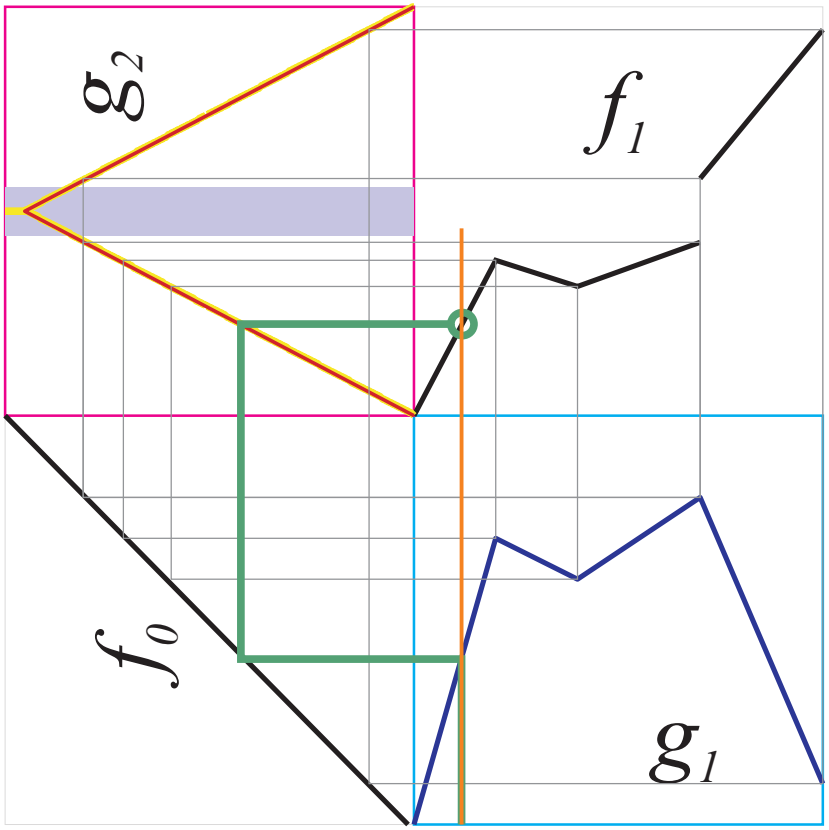


FIG. 20: **Quadweb illustration of the commutation operator** For ease of illustration, we choose f_0 as the identify map, and plot it in the lower left corner. The orange vertical shows a chosen x coordinate, and the green path shows the graphical computation of $g_2^{-1} \circ f \circ g_1(x)$. The intersection in the upper right quad is a point that lies on $f_1 = \mathfrak{C}_{g_1}^{g_2} f_0$. Generally, you would need to repeat for a large number of x coordinates. However, since all maps are linear in this example, a few well chosen points is sufficient to describe the first application of the commutation operator.

As a final illustration of alternative uses of the quadweb, we use the diagram to provide a graphical description of the action of the commutation operator. If we graph an arbitrary f_0 in the lower left panel, we can graphically compute $f_1 = \mathfrak{C}_{g_1}^{g_2} f_0$ as follows: (1) Choose an arbitrary x coordinate in the right half of the quadweb. Move up until you reach g_1, then left until you reach f_0, up again to g_2, then to the right. Although we used g_1 and f_0 in the usual way of graphically applying a function, we note that we used graph g_2 by going from range Y to domain Y, equivalent to applying g_2^{-1}. Consequently, we graphically "compute" the path $g_2^{-1} \circ f \circ g_1(x)$. The resultant intersection of that line with the original x coordinate is a point on f_1. Repeating at a sufficiently dense set of x coordinates will allow a robust description of f_1. See Fig. 20 for an example of this technique.

12. APPENDIX: ON ENTROPY AND INVARIANCE

The form of the conjugacy between the full logistic map $r = 4$ and the full tent map $a = 2$, which is $h(x) = \frac{1}{2}[1 - \cos(\pi x)]$, has been known for a long time, dating to the work of Ulam and Von Neumann in the late 1940's [7], and as we have said, it is still used widely as a primary example of topological conjugacy.

However, in some crucial ways, it is a very unusual example to consider typical of conjugacies between pairs of maps in general, as we now know through the studies herein. The reason that it should be considered as an aytypical example of conjugacy is that it is actually a smooth conjugacy, or a diffeomorphism, which carries with it a stronger equivalence between the two dynamical systems than we should expect between two dynamical systems which are only continuously but not smoothly related. This one example in all of the textbooks leads to an oversimplified, meaning overly generalized, intuition of what conjugacies are like "in general" between any (a, r) conjugate pair. We believe that this leads to a misconstrued intuition as to what to expect from conjugate dynamical systems in general.

A common point of view when comparing dynamical systems and to measure the complexity of orbits of a dynamical is to consider its entropies, [8, 9, 10]. Given a measure space, $(X, \mathcal{B}, \mu)$ for a μ-measure preserving dynamical system g, the entropy of a partition Q is defined,

$$H(Q) = -\sum_{m=1}^{k} \mu(Q_m) \ln \mu(Q_m), \tag{75}$$

from which the measure theoretic entropy with respect to the partition Q is defined,

$$h(T, Q) = \lim_{N\to\infty} \frac{1}{N} H(\vee_{n=0}^{N} g^{-n}(Q)), \tag{76}$$

and the measure theoretic entropy, also known as the Kolmogorov-Sinai entropy, is defined,

$$h_{KS}(g) = \sup_{Q} h(T, Q), \tag{77}$$

where the supremum is taken over all partitions. This may be interpreted as the amount of information gained per unit time by "typical" orbits. A useful fact is that KS entropy is invariant with respect to isomorphisms [8], meaning the transformation between the two dynamical systems is a bijection, but need not be continuous. A useful identity due to Pesin [12],

$$h_{KS} = \sum_{\lambda_i > 0} \lambda_i, \tag{78}$$

states that KS-entropy is the sum of positive Lyapunov exponents. This implies that a diffeomorphism causes preservation of KS entropy, since diffeomorphisms preserves Lyapunov exponents.

The topological entropy, on the other hand, may be defined in terms of metric entropies with respect to a "maximal entropy measure," ([25], p. 176) from which it can be understood,

$$h_T \geq h_{KS}. \tag{79}$$

Topological entropy has an intepretation as the (unweighted) growth rate of the number of states of a dynamical systems, or for certain C^1 diffeomorphism on compact manifolds, as the growth rate of the number of periodic orbits of a dynamical system through [11],

$$h_T(g) = \limsup_{N\to\infty} \frac{ln\ N_n(g)}{n}, \tag{80}$$

where $N_n(g)$ is the number of period-n orbits of g.

Topological entropy is relevant for our discussions concerning comparing dynamical systems, since h_T is invariant between two dynamical systems which are conjugate. However, the converse is not true. If the topological entropy of two dynamical systems is equal, then it does not follow that the two dynamical systems are conjugate. The example shown in Fig. 15 makes this clear. Each contains maximal invariant chaotic sets with identical shift symbolic representations, but for one, that shift is not the stable invariant set. Thus, if the goal is to compare two dynamical systems, then entropy alone is not sufficient, although of course it is a useful and suggestive quantity.

---

[1] Les m'ethods nouvelles de la m'ecanique c'eleste, Paris : Gauthier-Villars, 1892 vol. I , 1893 vol. II, 1899 vol. III (New methods of celestial mechanics, American Institute of Physics, 1993)

[2] Of course, given one dynamical system g_1, one can always construct a second dynamical by choosing an appropriate function $h(x)$ as an arbitrary change of coordinates, and simply constructing a new function $g_2 \equiv h \circ g_1 \circ h^{-1}$. However, refer to the harder question of finding h, if it exists, when g_1 and g_2 is given.

[3] E. Bollt, T.Stanford, Y-C. Lai, K. Zyczkowski, "What Symbolic Dynamics Do We Get With A Misplaced Partition? On the Validity of Threshold Crossings Analysis of Chaotic Time-Series," Physica D. 154 3-4 259-286 (2001).

[4] P. Grassberger, H. Kantz, Phys. Lett. A **113** 235 (1985); P. Grassberger, H. Kantz, U. Moenig, J. Phys. A: Math Gen. **22** 5217 (1989).

[5] P. Collet, J.P. Eckmann, *Iterated Maps on the Interval as Dynamical System,* (Birkhauser, 1980).

[6] Robert L. Daveney, *An Introduction to Chaotic Dynamical Systems, 2nd Edition,* (Westview Pr, 2nd edition, January 2003)

[7] S.M. Ulam and J. von Neumann, "On combination of stochastic and deterministic processes " (abstract). Bull. Amer. Math. Soc 53 (1947) 1120.

[8] Edward Ottt, *Chaos in Dynamical Systems, 2nd Edition*, (Cambridge U. Press, 2002).

[9] John Guckenheimer, Philip Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Applied Mathematical Sciences Vol. 42)*, (Springer, 2002).

[10] Alwyn Scott, *Encyclopedia of Nonlinear Science,* (Taylor and Francis, 2004).

[11] C. Robinson, *Dynamical Systems; Stability, symbolic dynamics, and chaos, 2nd ed.,* (CRC Press, 1999).

[12] J.P. Eckmann, D. Ruelle, "Ergodic Theory of chaos and strange attractors," Reviews of Modern Physics, 57: 617-656 (1985).

[13] B.P. Kitchens, *Symbolic Dynamics, One-sided, Two-sided and Countable State Markov Shifts,* Springer (New York, 1998).

[14] James Munkres, *Topology (2nd Edition),* (Prentice Hall, 2nd edition, 1999).

[15] J. Hale, *Ordinary Differential Equations,* (Robert E. Kreiger Publishing CO., Fld. 1980).

[16] A.N. Kolmogorov, S.V. Fomin, *Introductory Real Analysis*, (Dover Publications, New York, 1975).

[17] P.R. Halmos, *Measure Theory,* (de. Van Nostrand, New York, 1950).

[18] G. de Rham, "Sur quelques courbes définies par des équations fonctionelles," Rend. Sem. Mat. Torino 16, 101-113 (1956).

[19] L. Berg, M. Krüppel, "de Rham's Singular Function and Related Functions," Zeitschrift für Analysis und ihre Anwendungen 19 1 227-237 (2000).

[20] K.M. Brucks, M. Misiurewicz, and C. Tresser, Monotonicity properties of family of trapezoidal maps, Comm. Math. Phys. 137, 1 (1991).

[21] K.Zyczkowski, E.Bollt, "On the entropy devil's staircase in a family of gap-tent maps," Physica D 132 3 392-410 (1999).

[22] P. Grassberger and H. Kantz, "Generating partitions for the dissipative Hénon map," *Physics Letters A*, 113, pp 235-238, 1985.

[23] M. Misiurewicz, E. Visinescu, "Kneading sequences of skew tent maps. Annales de l'institut Henri Poincar (B) Probabilits et Statistiques," 27 1 125-140 (1991).

[24] P. Billingsley, *Probability and Measure, 3rd Edition,* (Wiley-Interscience, 3rd edition, 1995).

[25] Anatole Katok, Boris Hasselblatt,, *Introduction to the Modern Theory of Dynamical Systems (Encyclopedia of Mathematics and its Applications),* (Cambridge University Press, 1996).

[26] Dr. Seuss, *One fish, two fish, red fish, blue fish*, (Random House, New York, 1960.) [We recognize that literally the whole graph $\hat{g}_{2i}^{-1}$ would be a *zed*, with the portion extending beyond g_2 being the "hair" upon the head, we consider our definition to be only a minor abuse of notation.]

[27] Guo-Cheng Yuan, James A. Yorke, Thomas L. Carroll, Edward Ott, and Louis M. Pecora, "Testing Whether Two Chaotic One Dimensional Processes Are Dynamically Identical," Phys. Rev. Lett. **85** 20 4265-4268 (2000).

[28] The major step, of choosing and matching partitions of the two dynamical systems is somewhat easier in one-dimensional phases spaces, with interesting and important consequences if for some reason the generating partitions are not matched, as studied in our previous work, [3]. It may often be useful for complexity reduction of the models to deliberately not choose generating partitions, as exemplified by the examples in Figs. 11 and 12. In multivariate dynamical system, choosing the partitions to compare will be a challenging part of the problem in future work. For example the generating partition for certain diffeomorphisms of the plane are conjectured to coincide to a curve connecting "primary" heteroclinic tangencies [22], but again there is useful information to be gained by matching without regard to the generating partition [3].

[29] *Cover* — with each $y \in Y$ we associate a point $(y, g_2(y)) \in Y \times Y$. the collection of all such points is called the *graph* of g_2. A point $(y, g_2(y))$ on that graph is covered by the inverse function $\hat{g}_{2i}^{-1}$ if there is a $p \in Y$ such that $(\hat{g}_{2i}^{-1}(p), p) \equiv (y, g_2(y))$.

[30] One may easily create a measuring definition that is symmetric, for instance, by defining $\bar{\delta} = \delta(g_1, g_2) + \delta(g_2, g_1)$.

[31] Ideally, we would want equality to imply the converse for each of these four statements (i.e. $\lambda_O = 0 \Longrightarrow f$ is onto). However, because our approach to defining these defects is measure-based, the strongest condition we should expect is that we would satisfy these properties *almost everywhere.*

[32] By "folding," we mean to measure many-to-oneness, quantifying the amount of the range which has multiple pre-images.

[33] By *atomic part*, we mean to use the Lebesgue decomposition of a function [16] into its continuous part and atomic part, $f(x) = c(x) + a(x)$. The theorem applies to functions of bounded variation, and sometimes f will not satisfy this hypothesis. However, since we do not require an actual decomposition, we view our verbage as a minor abuse of notation. See [17] for a further discussion of this general decomposition.