# Physics 449 Thesis
# Topological Entropy as a Practical Tool for the Identification and Characterization of Chaotic Systems

Steven Gribble

April 12, 1995

*Department of Physics, The University of British Columbia*
*Vancouver, British Columbia, V6T 1Z1*

Supervisor - Dr. Kristin Schleich

# Contents

# List of Figures

## Abstract

The standard definition of chaos, attributed to J. A. Yorke, requires a dynamical system to have at least one positive Lyapunov exponent and to exhibit topological ergodicity in order to be considered chaotic. Lyapunov exponents, however, are an unsuitable tool for characterizing sensitive dependence on initial conditions when applied to systems studied in general relativity, because of their lack of invariance under time reparameterization. An alternate definition of chaos is presented - a system is chaotic if it has nonzero topological entropy. Various standard systems are explored, and both of the old and new definitions of chaos are applied in order to demonstrate their equivalence. In the course of the exploration, a new method for calculating Lyapunov exponents based on geodesic deviation is presented. Finally, an algorithm for the numerical appromation of topological entropy is proposed. This numerical scheme is applied to both the one dimensional logistic equation and the two dimensional Hénon map; difficulties encountered are discussed, and possible remedies proposed.

# 1 Introduction

Knowledge of the existence of ergodic and chaotic behaviour is not a new phenomenon. Nicole Oresme, in his treatise *Tractatus de commensurabilitate vel incommensurabilitate motuum cell[7]* (dated 1351), examined the motion of two bodies on a circle having constant but incommensurable velocities, and concluded that:

> No sector of a circle is so small that two such mobiles could not conjunct in it at some future time, and could not have conjuncted in it sometime [in the past].

This is one of the earliest recorded observations of ergodic behaviour. Recently, chaotic systems have been under intense scrutiny, but surprisingly there is still a lack of consensus on a rigorous definition of chaos. Informal definitions abound - an example compendium by Gleick[5] is:

- "The complicated, aperiodic, attracting orbits of certain dynamic systems" (Philip Holmes).
- "A kind of order without periodicity." (Hao Bai-Lin).
- "Dynamics with positive, but finite, metric entropy." (James Crutchfield).
- "Dynamics freed at last from the shackles of order and predictability." (Joseph Ford).

A more formal definition of chaos (due to Yorke) is presented in [19].

**Definition 1 (Chaos)** *Given a flow $\phi(t,x)$ and a compact set $\Lambda \subset \mathbb{R}^n$ invariant under $\phi(t,x)$, then $\phi(t,x)$ is said to be chaotic (w.r.t. $\Lambda$) if:*

1. *$\phi(t,x)$ has sensitive dependence on initial conditions on $\Lambda$.*
2. *$\phi(t,x)$ is topologically transitive on $\Lambda$.*
3. *The periodic orbits of $\phi(t,x)$ are dense in $\Lambda$.*

The third condition is often omitted for general systems; its use is as detector of the onset of turbulence.

Sensitive dependence on initial conditions is usually detected through the existence of at least one positive Lyapunov exponent, while the presence of topological transitivity and dense orbits is equivalent to topological ergodicity. (A flow $\phi(t,x)$ is topologically transitive with respect to a closed invariant set $A$ if, for any two open sets $U, V \subset A$, $\exists t \in \mathbb{R}$ such that $\phi(t,u) \cap V \neq \emptyset$.) We may therefore paraphrase the standard definition 1 and state that a system is chaotic if it has at least one positive Lyapunov exponent, and is topologically ergodic.

One can imagine that two neighboring trajectories of a chaotic system are being driven apart by a positive Lyapunov exponent (although this misleadingly gives Lyapunov exponents an active role in the evolution of a system - they are simply a measurable property of a system). The contrasting element, ergodicity, both bounds the trajectories to an invariant region, and requires that trajectories densely cover that invariant region. It is the conflict created by the presence of both of these qualities that results in chaotic behavior.

The Yorke definition of chaos (which we will take as the standard definition) is one of many different definitions in active use by researchers. All definitions attempt to confirm the presence of the two salient features identified by the standard definition (ergodicity and sensitive dependence on initial conditions). Tests which attempt to identify strange attractors, demonstrate the existence of a Cantor set, or compute the fractal dimension of an attracting set can all be shown to be equivalent to the standard Yorke definition; the presence of these secondary features is a consequence of ergodicity and sensitive dependence on initial conditions.

It should be noted that the standard definition is particularly suitable for application to Hamiltonian flows (which are relevant to a large number of physical systems). Indeed, it is extremely tempting to use this definition in the study of chaotic systems in general relativity; however, one encounters difficulties in doing so. The standard definition relies on Lyapunov exponents as an indicator of sensitive dependence on initial conditions; however, Lyapunov exponents are not invariant under time reparameterization (section 2.3). By reparameterizing time, one simply alters the coordinate system in general relativity; such coordinate changes preserve the physics of the spacetime. Unfortunately, by performing such a reparameterization it is possible to change a positive Lyapunov exponent to have a zero value. Sensitive dependence therefore is not well characterized by Lyapunov exponents in general relativity, which leaves us searching for a more suitable definition of chaos that is invariant under time reparameterization.

This thesis explores alternative definitions of chaos that can be safely applied to a Hamiltonian system whose physics is invariant under time reparameterization, such as those which occur in the study of general relativity. The definition of chaos that we will consider is similar to that of James Crutchfield. We define a system to be chaotic if that system has non-zero (but finite) topological entropy. Topological entropy (discussed in section 3) essentially is a quantitative description of the 'rate of mixing' of a dynamical system. This definition has been considered in the past, but the seemingly intractable nature of analytic or numerical calculations of topological entropy has always discouraged the used of this definition, and encouraged the reliance on the more practical (but occasionally insufficient) tool of Lyapunov exponents.

In this thesis, a number of example systems will be discussed and the previously mentioned properties will be calculated, and the equivalence of the standard definition (1) and our alternate definition based on topological entropy will be demonstrated. In the course of these analyses, a new method for the calculation of Lyapunov exponents based on geodesic deviation will be introduced. Also, a paradigm for the numerical approximation of topological entropy suitable for application to systems studied in general relativity will be presented.

Throughout this thesis, we define a dynamical system in terms of a flow on a manifold. Most standard theorems presented in this thesis are proven for differentiable flows; ie. diffeomorphisms, which are smooth, 1-to-1 onto maps with smooth inverses. It should be noted, however, that many (but not all) of these theorems can be generalized to the case of measurable flows. Given a flow $T : \mathcal{X} \to \mathcal{X}$, we say that $T$ is a measurable flow if and only if $T^*\mu(A) = \mu(T^{-1}(A))$. That is, if $T(A)$ is measurable, then so is $T^{-1}(A)$. (Refer to section 2.1 for a discussion of measure theory.)

# 2 Conventional Characterizations of Chaos

In this section of the thesis, we will explore in detail the two requirements a system must satisfy according definition 1 in order to be deemed chaotic - namely Lyapunov exponents and ergodicity.

## 2.1 Ergodicity

The standard definition of chaos (definition 1) includes the requirement of topological ergodicity. Ergodicity is related to how a system folds back on itself and mixes. If a system is ergodic, then almost every flow will densely cover the phase space of the system. There are many different formal definitions of ergodicity, but it is important to note that all of these definitions are completely equivalent.

Definitions of ergodicity are usually given in terms of invariant probability measures. A measure $\mu$ can be considered [1] to be a function that, given a region $A$ in a manifold $\mathcal{M}$, returns a real number:

$$\mu(A) : \mathcal{M} \to \mathbb{R}, \text{with } \mu(A) = \int_A d\mu. \tag{2}$$

$\mu$ is a probability measure if $\mu \geq 0$ and $\mu(\mathcal{D}) = 1$ where $\mathcal{D}$ is the entire space in question. $\mu$ is an invariant of the map $f : \mathcal{M} \to \mathcal{M}$ if:

$$\mu(A) = \mu(f^t(A)) \qquad \text{(for all t, and for } A \in \mathcal{M}). \tag{3}$$

### 2.1.1 Ergodicity and Statistical Mechanics

Consideration of ergodicity first arose in the context of statistical mechanics in the form of the ergodic hypothesis put forward by Boltzmann and Gibbs. Given a system of N particles, one can represent the state of the system by a single point $\mathbf{x}$ in 6N dimensional phase space. If the energy $E$ of the system is constant, then the motion of the system is restricted to the $6N-1$ dimensional integral manifold $\mathcal{M}$ specified by $H(\mathbf{x}) = E$. Given some continuous function $\phi(\mathbf{x})$ that represents the instantaneous value of some property of the system, the time average value of that property along the path $x(t)$ is given by:

$$\bar{\phi} = \lim_{t\to\infty} \frac{1}{t} \int_{\tau=0}^{\tau=t} \phi(\mathbf{x}(\tau)) d\tau. \tag{4}$$

Equation (4) is impractical in nature, because if the dynamics of the system are complex, then $\mathbf{x}(\tau)$ may be difficult to evaluate. It would be much simpler to use the space average of the system as the macroscopic value of the property. This space average is given by:

$$\langle\phi\rangle_{space} = \int_{\mathcal{M}} \phi(\mathbf{x}) d\mu \tag{5}$$

where $\mu$ is an appropriate probability measure. The ergodic hypothesis then puts forth the claim that:

$$\langle\phi\rangle_{space} = \bar{\phi} \tag{6}$$

[1] More generally, $\mu$ can be thought of as a continuous linear functional. Given a function $\phi : \mathcal{M} \to \mathbb{R}$, $\mu$ is defined relative to some density $h(x)$ by:

$$\mu(\phi) \equiv \int_{\mathcal{M}} \phi(x) h(x) dx \tag{1}$$

where we typically use the notation $h(x)dx = d\mu$.

In fact, we can take (6) to be another definition of ergodicity.

**Definition 2 (Ergodicity)** *A system is ergodic with respect to a measure $\mu$ if and only if the space average $\langle\phi\rangle_{space}$ is equal to the time average $\bar{\phi}$. Such a $\mu$ is said to be an ergodic measure.*

Definition 2 is not the only definition of ergodicity. Another intuitive definition is related to the notion of recurrence. The Poincaré Recurrence theorem states:

**Theorem 1 (Poincaré Recurrence)** *Consider a diffeomorphism $f^t$ on a manifold $\mathcal{M}$, and an invariant measure $\mu$ such that $\mu(\mathcal{M}) < \infty$. For any region $A \in \mathcal{M}$ such that $\mu(A) > 0$, $\exists\, x \in A$ such that $f^n(x) \in A$ for some finite $n \in \mathbb{Z}$. Furthermore, this recurrence occurs for all points in $A$, except possibly for a set of measure zero. Furthermore, this recurrence occurs an infinite number of times.*

Poincaré recurrence guarantees that a trajectory of a system passing through (almost any) point $\mathbf{c}$ will eventually pass through an arbitrarily small neighborhood of $\mathbf{c}$, and in fact will do so an infinite number of times. If the system is ergodic, a much stronger statement can be made, namely that given a trajectory $\mathbf{x}(t)$ passing through a point $\mathbf{c}$, that trajectory will pass arbitrarily close to any other point of the system. This is true for all points $\mathbf{c}$ within the system, except possibly for a set of measure zero. We can use this fact to give another equivalent definition of ergodicity:

**Definition 3 (Ergodicity)** *If $T : \mathcal{M} \to \mathcal{M}$ is an ergodic mapping on a closed manifold $\mathcal{M}$, then $\exists A \subset \mathcal{M}$ such that $A, TA, \ldots, T^nA$ are pairwise disjoint and cover $\mathcal{M}$ up to set of measure less than $\epsilon$, for any $n > 0$ and $\epsilon > 0$.*

The following alternate definition of ergodicity is cited in [14]:

**Definition 4 (Ergodicity)** *Given the mapping $f : \mathcal{M} \to \mathcal{M}$, any invariant set $A$ of the mapping $f$, and a measure $\mu$, then $f$ is ergodic with respect to $\mu$ if $\mu(A) = 0$ or 1.*

A related definition of ergodicity (cited in [12]) can be given as:

**Definition 5 (Ergodicity)** *Given the mapping $f : \mathcal{M} \to \mathcal{M}$ which preserves the probability measure $\mu$, then $f$ is ergodic if and only if every invariant, measurable function $g = g \circ f$ on $\mathcal{M}$ is constant almost everywhere.*

It should be stressed that these definitions of ergodicity are all completely equivalent. In fact, Peterson[12] explicitly proves the equivalence of these defintions; we may therefore use any of them in order to demonstrate the ergodicity of a given dynamical system.

### 2.1.2 Features of Ergodicity

In contrast to Lyapunov exponents, ergodicity is a global, time independent quality of a system. All four given definitions of ergodicity stress this global nature: the entire domain of a system needs to be taken into account in order to determine whether or not the system is ergodic. The system must also be examined over all time, not just at one instantaneous moment, resulting in an immunity to rescalings and dilations of time.

It should be noted that at least one ergodic measure can be found for any differentiable flow on a compact manifold $\mathcal{M}$ [13]. For example, the phase space of a linear oscillator can be though of as being composed of an infinite number of integral manifolds, each of which corresponds to a surface of constant energy. If the system lies on a particular energy surface $\mathcal{E}$, then the system will densely cover that surface as it evolves over time. An ergodic measure can be found corresponding to this dense covering, but it will only span the integral manifold, and will be zero everywhere else. We are only interested in system for which the ergodic measure spans the entire (bounded) domain of the system, and not just an integral manifold or some other subregion of the system. For this reason, we may automatically exclude any integrable system from the possibility of being chaotic. (However, we are free to consider an non-integrable submanifold of that system as an entirely new system which may exhibit chaoticity.) We define such non-integrable ergodicity as *topological ergodicity*.

## 2.2 Sensitive Dependence on Initial Conditions and Lyapunov Exponents

Nearby initial states of a system that exhibits sensitive dependence on initial conditions can be observed to diverge from each other for future times. Consider two initially nearby trajectories of a dynamical system; if these two trajectories rapidly diverge from each other after a short amount of time, then an experimenter needs to know the initial position with extremely high precision in order to predict the final state of the system. The Lyapunov exponents of a system are one measure of the degree of sensitivity of such a system. If the divergence of the two nearby trajectories is exponential in time (with $\lambda$ as the constant factor in the exponent), then $\lambda$ is a Lyapunov exponent of the system. In an $N$ dimensional system, there will be up to $N$ different $\lambda_i$.

There are a number of definitions of Lyapunov exponents which differ from each other based on the class of system that they characterize. However, the overall concepts behind the $\lambda_i$ remain the same. It is therefore sufficient to present the theory behind the calculation of the Lyapunov exponent of a 1-d difference equation in full detail, and only summarize the final results for other classes of systems.

### 2.2.1 1-d Difference Equations

Jackson[9] gives an excellent presentation of the physicists' definition Lyapunov exponents and 1-d difference equations, which is summarized here. Consider the difference equation $x_{n+1} = f(x_n)$. Let $x_0, y_0$ be two nearby initial points in the space $\mathbb{R}^1$, with $x_n = f^n(x_0)$ and $y_n = f^n(y_0)$. If these points separate exponentially in $n$, then $|y_n - x_n| = Ae^{\lambda n}$ where $A = |y_0 - x_0|$. For large $n$, $\frac{1}{n}\ln|y_n - x_n| \to \lambda$. Of course, if the motion is in a bounded region, separation cannot occur for large $n$ unless $x_0, y_0$ are correspondingly close. We therefore must use:

$$\lambda = \lim_{n\to\infty}\frac{1}{n} \qquad \lim_{|x_0-y_0|\to 0} \ln\left|\frac{x_n - y_n}{x_0 - y_0}\right| = \lim_{n\to\infty}\frac{1}{n}\ln\left|\frac{df^n(x_0)}{dx}\right|. \tag{7}$$

(Refer to figure 1.) However, the chain rule of differentiation tells us that:

$$\frac{df^n(x_0)}{dx} = \frac{df(x_{n-1})}{dx_{n-1}}\frac{df^{n-1}(x_0)}{dx} = \cdots = \prod_{i=0}^{n-1}\frac{df(x_i)}{dx_i} \tag{8}$$

Therefore, the Lyapunov exponent of the difference equation is given by:

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{df(x_i)}{dx_i} \right|. \tag{9}$$

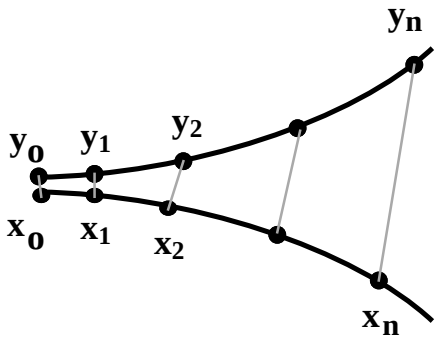



Figure 1: Lyapunov exponents and a 1-d difference equation

Finally, if a probability distribution $P(x)$ of the asymptotic ($n \to \infty$) set of points can be defined, then an average Lyapunov exponent $\bar{\lambda}$ over all initial states can be defined as:

$$\bar{\lambda} = \int \ln \left| \frac{df}{dx} \right| P(x) dx. \tag{10}$$

### 2.2.2 Lyapunov Exponents and Differentiable Dynamics

If we consider an n-dimensional system, we expect to be able to identify $n$ different Lyapunov exponents $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. From Ruelle[14], we can consider the case in which $\mathbf{f}$ is a differentiable map, giving us the time evolution

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n)). \tag{11}$$

We then label the Jacobian $D_x f = (\partial f_i / \partial x_j)$, yielding:

$$T_x^n = (D_{f^{n-1}(x)} f)(D_{f^{n-2}(x)} f) \cdots (D_{f(x)} f)(D_x f) = D_x f^n \tag{12}$$

where $f^n = f \circ f \circ \cdots \circ f$ ($n$ times). Defining

$$\Lambda_x = \lim_{n\to\infty} \left( T_x^{n^*} T_x^n \right)^{\frac{1}{2n}} \tag{13}$$

the Lyapunov exponents $\lambda_i$ of the differentiable map $\mathbf{f}$ are the logarithms of the eigenvalues of $\Lambda_x$. Note that the rate of growth of an n-dimensional volume element is given by $|\det(\partial f_i / \partial x_j)|$, from which we may also conclude that:

$$\ln |\det(\partial f_i / \partial x_j)| = \lambda_1 + \lambda_2 + \ldots + \lambda_n \tag{14}$$

### 2.2.3 Continuous Maps and Numerical Techniques

In contrast to difference equations $x_{n+1} = f(x_n)$, a 1-d continuous map has the form $x(t) = f^t(x(0))$. This continuous map can easily be discretized in segments of length $\Delta t$, resulting in the difference equation $x_n = f^{n\Delta t}(x_0)$. Equations (9) and (10) can then be directly applied to this difference equation. A similar technique can be used to analyze n-dimensional continuous maps using a variational approach.

If one were to pick a pair of vectors at random and evolve them through map (11), then for almost all pairs the vectors would diverge exponentially with the rate of growth of separation distance given by $\lambda_1$ (the largest Lyapunov exponent). It is reasonably simple to implement a numerical simulation of the mapping in order to track two nearby initial points and estimate $\lambda_1$. In fact, one can simply look at the variational equation

$$\frac{d\mathbf{z}(t)}{dt} = \left(\frac{\partial f_i}{\partial x_j}\right)_{\mathbf{x}(t)} \mathbf{z}(t) \tag{15}$$

where $\mathbf{z}(t)$ is the perturbation of the path $\mathbf{x}(t)$. In this case, we know that $|\mathbf{z}(t)| \to A\exp(\lambda_1 t)$, or

$$\lambda_1(x) = \lim_{t\to\infty} \frac{1}{t} \ln|\mathbf{z}(t)|. \tag{16}$$

A straightforward numerical implementation of (16) will result in an accurate approximation of $\lambda_1$. Refer to Jackson[10] for a discussion on how to approximate the remaining $\lambda_i$.

## 2.3 Features of Lyapunov Exponents

There are two important features of Lyapunov exponents that limit their usefulness as a practical tool for characterizing chaotic systems. Firstly, Lyapunov exponents are *local* in nature. Lyapunov exponents of a system are defined relative to a single point or neighborhood of the system - it is not necessarily true that the Lyapunov exponents are constant throughout the system. This locality introduces some difficulties when using them to classify chaotic systems, and in part explains why ergodicity is also a required characteristic of chaos. In order to simulate a global analysis, the Lyapunov exponents of a system can be calculated for points sampled from many different regions, hopefully identifying any potentially non-chaotic regions. However, this analysis is inherently flawed, because even if a large number of points are sampled, there are infinitely more points that are not.

### 2.3.1 A More Rigorous Definition of Lyapunov Exponents

The disadvantages of the local nature of Lyapunov exponents can be somewhat remedied by adopting a more rigorous definition of Lyapunov exponents relative to an ergodic measure. Pollicott [13] gives the following theorem by Oseledec:

**Theorem 2 (Lyapunov Exponents)** *Let $f : \mathcal{M} \to \mathcal{M}$ be a $C^1$ diffeomorphism of a compact manifold of dimension $n$, and let $\mu$ be an ergodic measure. Then there exist:*

1. *real numbers $\lambda_1 > \ldots > \lambda_k (k \leq n)$*

2. *positive integers $n_1, \ldots, n_k \in \mathbb{Z}^+$ such that $n_1 + \ldots + n_k = n$*

*3. a measurable splitting* $T_x\mathcal{M} = E_x^1 \oplus \ldots \oplus E_x^k$, *with* $dim(E_x^i) = n_i$ *and* $D_x f(E_x^i) = E_{fx}^i$, *such that*

$$\lim_{n\to\infty} \frac{1}{n} \ln \|D_x f^n(v)\| = \lambda_l \text{ for all } x \in \mathcal{M} \text{ up to a set of measure zero.} \tag{17}$$

*whenever* $v \in (E_x^1 \oplus \ldots \oplus E_x^l)$, *but* $v \notin (E_x^1 \oplus \ldots \oplus E_x^{l-1})$.

If this definition is used (whose fundamental difference from previous definitions is the assertion that there is an ergodic measure $\mu$ on the manifold), then the behaviour becomes global up to a set of measure zero. Of course, as it was previously mentioned, an ergodic measure can be found for any time evolution on a compact manifold $\mathcal{M}$, but this measure may not have support over the entire manifold; the behaviour only becomes global on the region over which the measure has support. For this reason, the Oseledec theorem only helps us for topologically ergodic measures.

### 2.3.2 Time-Dependancy of Lyapunov Exponents

Another feature of Lyapunov exponents is that their definition is time-dependent. When considering systems arising in general relativity, it becomes apparent that this time-dependency is a serious drawback. Time is no longer an absolute quantity, and it may be radically dilated or rescaled by a coordinate change. In such a situation, one must be extremely careful about how a positive Lyapunov exponent should be interpreted. The following example (provided by Schleich[16]) clearly demonstrates this fact.

Consider the example of the de Sitter spacetime (written with spatially flat hypersurfaces). A common form of the metric is given by:

$$ds^2 = -dt^2 + a^2(t)(dx^2 + dy^2 + dz^2). \tag{18}$$

Coupling this metric to the cosmological constant $\Lambda$ yields the following Einstein equations:

$$3(\frac{1}{a}\frac{da}{dt})^2 = \Lambda \tag{19}$$

with two disjoint, independent solutions given by:

$$a(t) = a_0 e^{\pm Ht} \qquad (\text{with } \ 0 < a_0 < \infty) \tag{20}$$

Here, the positive sign indicates an expanding solution, while the negative sign corresponds to a collapsing solution. (Both have $H^2 = \frac{\Lambda}{3}$.) Because of the independence of the solutions, the Lyapunov exponent is calculated separately for each solution.

For the positive case, the considering the evolution of the deviation between two infinitesimally neighbouring solutions $\xi = a_1(t) - a_2(t)$ yields

$$\frac{6H}{a_0}(\frac{d\xi}{dt} - H\xi) = 0 \tag{21}$$

which has the solution $\xi = \xi_0 e^{Ht}$. Trivially, we see $\lambda_+ = H$ for this case. Similarly, we find $\lambda_- = -H$ for the collapsing solution.

Now we note that the theory of general relativity is invariant under diffeomorphisms; in particular, changes of coordinates affect the form of quantities such as the metric, but not their physical meaning or interpretation.

Time reparameterizations therefore leave the physics of a given spacetime unchanged. It is now possible to demonstrate that the Lyapunov exponents calculated for the dynamics of a spactime are changed under such a time reparameterization. Consider the reparameterization of the de Sitter spacetime by $\tau = e^{Ht}$. The metric in the new coordinate system is given by:

$$ds^2 = -\frac{1}{H^2\tau^2}d\tau^2 + a^2(\tau)(dx^2 + dy^2 + dz^2). \tag{22}$$

The dynamics of this metric are now given by:

$$3(\frac{H\tau}{a}\frac{da}{d\tau})^2 = \Lambda. \tag{23}$$

The two, independent solutions are

$$a(\tau) = \pm a_0\tau. \tag{24}$$

Now, the equation for infinitesimal deviation of the expanding solution is given by

$$\frac{6H^2}{a_0}(\frac{d\xi}{d\tau} - \frac{1}{\tau}\xi) = 0, \tag{25}$$

having a solution $\xi = \xi_0\tau$ from which we immediately conclude that $\lambda_+ = 0$. The collapsing solution similarly yields $\lambda_- = 0$.

We have therefore demonstrated that, through the judicious choice of a time reparameterization, it is possible to scale Lyapunov exponents from a positive or negative value to zero, and vice-versa. (We could have done even better by scaling a positive solution to a negative solution.) We are forced to conclude that Lyapunov exponents, by themselves, are not sufficient to indicate sensitive dependence to initial conditions when applied to general relativity.

# 3 An Alternate Characterization of Chaos

Given the fact that Lyapunov exponents are unreliable tools to use in characterizing potentially chaotic systems in general relativity, we now turn our attention to topological entropy as a possible replacement. The intended goal of doing so is to be able to find a characterization of sensitive dependence on initial conditions that is independent of Lyapunov exponents.

Much in the same way that the concept of ergodicity can be approached from various different perspectives, topological entropy can also be understood in a number of different ways. Intuitively, topological entropy describes the 'rate of mixing' of a dynamical system. As such, it can immediately be seen to have a relationship to both Lyapunov exponents, through the dependency on rate, and to ergodicity, because of the association with mixing. Alternatively, topological entropy can be thought of as an information theoretic property of a system. If an experimenter can only measure the state of a system to a finite degree of accuracy, then topological entropy measures the rate at which the experimentor can gain more information about the system as it evolves over time.

### 3.1 The Calculation of Topological Entropy

(Jackson[10]) Consider a compact space $X$ and a continuous map $f$ that takes the space into itself,

$$f : X \to X \tag{26}$$

Denote by $A = \{a_i\}$ a cover of $X$ if

$$a_i \in X \, \forall \, i, \; X \subset \bigcup_i a_i. \tag{27}$$

This cover is a *partition* if $(a_i \cap a_j = \emptyset, \forall \, j, k)$. Denote by $N(A)$ the number of regions in the partition. We now define the product of two partitions $A = \{a_i\}$ and $B = \{b_i\}$ of a space $X$ by:

$$C = A \vee B \equiv \{c_i = a_j \cap b_k, \forall \, j, k\} \tag{28}$$

In general, $N(A)N(B) \geq N(A \vee B) \geq N(A), N(B)$. (Refer to figure 2 for an illustration.)

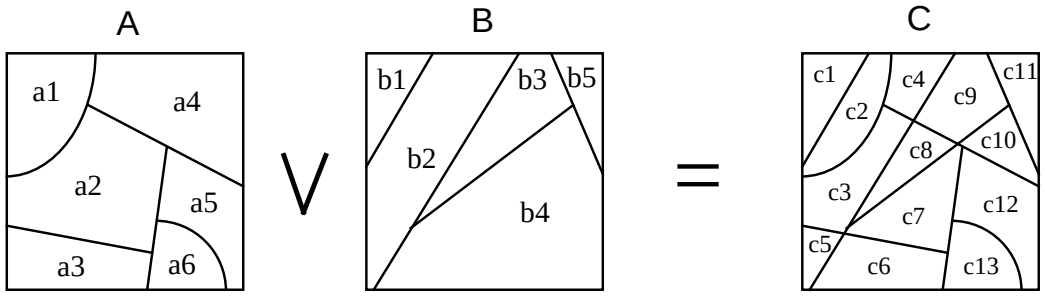



Figure 2: The product of two partitions

The map $f$ takes the partition $A$ into a new partition $f(A) = B = \{f(a_i)\}$. Since $f$ may be many-to-one instead of one-to-one, we should instead consider the inverse of the map: $f^{-1}(A) = B = \left\{f^{-1}(a_i)\right\}$. Define $A^n$ by:

$$A^n \equiv A \vee f^{-1}(A) \vee \ldots \vee f^{-(n-1)}(A) \equiv \bigvee_{k=0}^{n-1} f^{-k}(A). \tag{29}$$

The T-entropy $H(A)$ of a partition $A$ is:

$$H(A) = \ln N(A). \tag{30}$$

The T-entropy $h(f, A)$ of a mapping $f$ with respect to a partition $A$ is:

$$h(f, A) = \lim_{n \to \infty} \frac{1}{n} H \left( \bigvee_{k=0}^{n-1} f^{-k}(A) \right). \tag{31}$$

Finally, we define the topological entropy $h(f)$ of a mapping $f$ as the supremum of (31) over all partitions A:

$$h(f) = \sup_A h(f, A) = \sup_A \lim_{n \to \infty} \frac{1}{n} H \left( \bigvee_{k=0}^{n-1} f^{-k}(A) \right). \tag{32}$$

It is obviously very impractical to calculate $h(f, A)$ for all partitions $A$ in order to evaluate the supremum in (32). However, this can be avoided by considering refinements of partitions. The partition $B = \{b_i\}$ is a refinement of $A = \{a_i\}$ if, for all $i$, $b_i \in a_j$ for some $j$. This is denoted by $B > A$. We can now define a refining sequence of partitions:

$$\{A_n | n = 1, 2, \ldots\} \text{ where } A_n < A_{n+1} \text{ and for all } B, \exists A_n > B. \tag{33}$$

Adler, Konheim, and McAndrew [2] have shown that if $\{C_n\}$ is a refining sequence, then:

$$h(f) = \lim_{n \to \infty} h(f, C_n). \tag{34}$$

This equation for $h(f)$ has far more practical value than (32).

Topological entropy can be similarly defined for a flow $f : A, t \to A$. In this case, Adler, Konheim, and McAndrew [2] have supposed that

$$h(f_t) = |t| \, h(f_1). \tag{35}$$

In this way, the discrete mappings that we have been considering can be thought of as a flow over a unit time period.

## 3.2 Features of Topological Entropy

In many senses, topological entropy provides a true measure of the chaoticity of a system. It is global in nature, and shares with ergodicity an immunity to rescalings of time. In order for a system to have non-zero topological entropy, the rate of mixing must be exponential, which is reminiscent of Lyapunov exponents; however, this exponentiality of mixing isn't relative to time, but rather to the number of discrete steps through which the system has evolved - it is because of this that topological entropy is immune to reparameterizations of time.

## 3.3 Current Numerical Methods

In order to place a lower bound on the topological entropy of a map experimenters often make use of the Pesin identity 36, which provides a way to relate the Lyapunov exponents of a map to *measureable entropy*. The following definition of easurable entropy is presented by Pollicott[13].

Assume that $f : \mathcal{M} \to \mathcal{M}$ is a $C^1$ diffeomorphism on a compact manifold. We want to define a map $h_{meas} : \mathcal{M}_{inv} \to \mathbb{R}^+$ which associates to each $f$-invariant probability measure $\mu$ a positive real number $h_{meas}(u) \geq 0$, as follows:

1. If $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}^+$ is the metric on $\mathcal{M}$ (derived from the Riemannian metric), then for any $n \geq 1$ define a new metric by
$$d_n(x, y) = \max_{0 \leq i \leq n} d(f^i x, f^i y).$$
2. For all $0 < \delta < 1$, $n \geq 1$, $\epsilon > 0$ we call a finite set $K \subseteq M$ an $(n, \epsilon, \delta)$-covering set if the union $\cup_{x \in K} D_\epsilon(x; d_n)$ of the $\epsilon$-balls $D_\epsilon(x; d_n) = \{y \in M | d_n(x, y) < \epsilon\}$ centered at points $x \in K$ has $\mu$-measure greater than $\delta$.

3. Let $N(n,\epsilon;\delta)$ be the smallest possible cardinality of a $(n,\epsilon,\delta)$-covering set (i.e. $N(n,\epsilon;\delta) = \min\{$ card $K | K$ is a $(n,\epsilon;\delta)$-covering set $\}$.)

4. Finally, we define the measure theoretic entropy of the measure $\mu \in \mathcal{M}_{inv}$ as:

   **Definition 6 (Measurable Entropy)**

$$h_{meas}(\mu) = \lim_{\delta \to 1} \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \ln N(n,\epsilon;\delta) \geq 0$$

Pesin[11] has proven that, if $f$ is a $C^{1+\alpha}$ diffeomorphism of a compact manifold $\mathcal{M}$ that preserves a smooth ergodic measure $\mu$, then

$$h_{meas}(f) = \sum_{\lambda_i > 0} \lambda_i. \tag{36}$$

Not only does this Pesin identity provide a direct connection between Lyapunov exponents and measurable entropy, but it also has practical significance. In order to numerically approximate measurable entropy, the current practice is to numerically determine the Lyapunov exponents of a system, and then use the Pesin identity. We can further use the fact that $h_{meas} \leq$ topological entropy $h$ to use the Pesin identity as a generator of a lower bound on the topological entropy.

## 3.4 KS-entropy

A concept that is closely related to that of measurable entropy is Kolmogorov-Sinai-entropy, or KS-entropy. Given a continuous map $f$, an invariant region $X$, and an invariant probability measure $\mu(a)$, s.t.

$$\mu(a) = \mu(f^{-1}(a)) = \mu(f(a)) \qquad \text{for all } a \in X. \tag{37}$$

Assume further that $A = \{a_i\}$ is a partition, with $X = \bigcup_{i=1}^{N} a_i$, $\mu(X) = 1$, and $\mu(a_i \cap a_j) = 0 \; \forall \; i \neq j$. Then, define:

$$H_\mu(A) \equiv -\sum_{i=1}^{N} \mu(a_i) \ln \mu(a_i). \tag{38}$$

The KS-entropy of the mapping $f$ for the partition $A$ is given by:

$$h_\mu(f,A) = \lim_{n \to \infty} \frac{1}{n} H_\mu \left( A \vee f^{-1}A \vee \ldots \vee f^{-(n-1)}A \right). \tag{39}$$

The KS-entropy of the map is then given by

$$h_\mu(f) = \sup_A h_\mu(f,A). \tag{40}$$

As with topological entropy, it is for the most part infeasible to evaluate the supremum in equation (40). Fortunately, an idea similar to (34) can be applied, in that given a refining sequence $\{C_n\}$, we can again state:

$$h_\mu(f) = \lim_{n \to \infty} h_\mu(f, C_n). \tag{41}$$

It has been conjectured[2] that

$$h(f) = \sup_{\mu} h_{\mu}(f), \tag{42}$$

where the supremum is taken over all normalized Borel measures invariant through the map $f$. (Refer to appendix B for a discussion of Borel measures and measure spaces in general.) It is certainly the case, however, that

$$h(f) \geq h_{\mu}(f), \tag{43}$$

with equality if $\mu$ is taken as the Lebesgue measure and the map $f$ is area-preserving. In effect, this means that $h_{\mu}(f)$ is a more detailed measure than $h(f)$. Indeed, $h(f)$ may only be zero or positive - $h_{\mu}(f)$ can be negative, zero, or positive. In either case, we only consider mappings with positive entropy (T-entropy or KS-entropy) to be chaotic.

# 4 Chaotic and Non-Chaotic Systems

In this section of the thesis, we will consider a number of dynamical systems in order to build a repertoire of chaotic and non-chaotic behavior. These systems will be analyzed for their Lyapunov exponents, ergodicity, and where possible, topological entropy. Both analytical and numerical methods will be used in the analyses - by doing so, an intuitive feeling will hopefully be developed for what is involved in each type of analysis, as well as for the strengths and weaknesses of each of the methods.

## 4.1 One Dimensional Logistic Equation

A famous one-parameter system that exhibits both chaoticity and non-chaoticity for different values of the parameter is the one-dimensional logistic equation. The logistic equation dates back to ecologists' models of population dynamics. The one-dimensional logistic equation is:

$$x_{n+1} = f(x_n) = px_n(1 - x_n), \qquad 0 \leq p \leq 4. \tag{44}$$

The $px_n$ term represents the Malthusian influence of unlimited population growth in the absence of limiting forces, while the $(1 - x_n)$ term imposes a limiting factor on this growth term. Typically, we restrict our study to the interval $[0, 1]$. When $x_n$ approaches 1, $1 - x_n$ becomes small and drives the population back down towards 0.

It has been demonstrated that this map has an attractive fixed point for $p < 3$ and has an attractive periodic orbit (of period $2^n$) with $n \to \infty$ as $p \to 3.57\ldots$. At $p = 4$ (and most of the other values for which $p > 3.57\ldots$) the system exhibits chaoticity.

For the purposes of this section, we will look at the $p = 4$ case. Consider the homeomorphism of $[0, 1]$ into itself given by $\phi$ (Ruelle[14], p. 41):

$$\phi(x) = \frac{2}{\pi}\arcsin x^{\frac{1}{2}}. \tag{45}$$

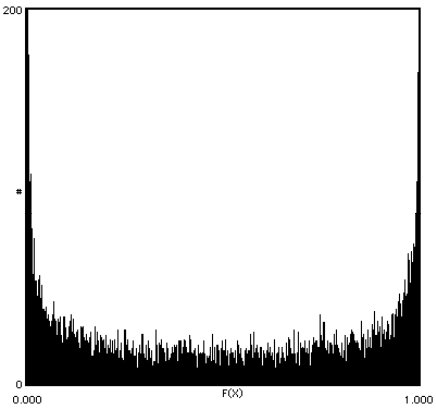
Then, we define

$$\tilde{f}(x) = \phi \circ f \circ \phi^{-1}(x) = \begin{cases} 2x & \text{if } x \in \left[0, \frac{1}{2}\right] \\ 2 - 2x & \text{if } x \in \left[\frac{1}{2}, 1\right] \end{cases} \tag{46}$$

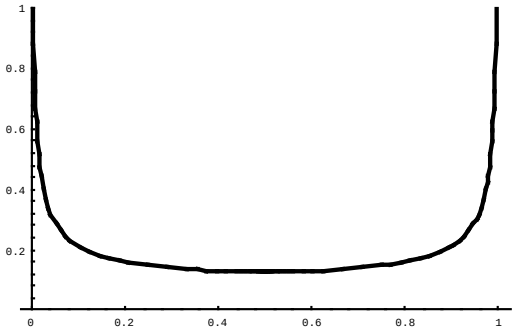
The map (46) is a tent map, which preserves the Lebesgue measure [2] $m$. Because of this, it can be seen that the logistic equation preserves the measure $\mu = \phi^{-1}m$, or:

$$d\mu = \frac{d\phi}{dx}dx = \frac{1}{\pi}\frac{dx}{[x(1-x)]^{\frac{1}{2}}} \tag{47}$$

$\mu$ is thus a continuous, invariant probability measure for the map $f$. Figure 3b illustrates the theoretical probability distribution and figure 3a provides a numerical probability distribution obtained by iterating the initial point .438293947 through (44) 100,000 times.




(a) Observed probability distribution (100,000 iterations)




(b) Theoretical probability distribution

Figure 3: The logistic equation probability distribution

#### 4.1.1 Lyapunov Exponent

We can use equation (10) to analytically calculate the average lyapunov exponent $\bar{\lambda}$ of the logistic equation (with $p = 4$), given the probability distribution (47). Thus, we have:

$$\begin{aligned} \bar{\lambda} &= \int_0^1 \ln\left|\frac{df(x)}{dx}\right| d\mu \\ &= \int_0^1 \ln|4-8x| \frac{dx}{\pi(x(1-x))^{\frac{1}{2}}} \\ &= \ln 2 \end{aligned} \tag{48}$$

It is also possible to numerically estimate $\bar{\lambda}$ through the use of equation (9). Using $\frac{df(x_i)}{dx_i} = (4-8x)$, a simple C program was written which simply iterated through (9) for $i = 0$ through $100,000$, and yielded the approximate value $\lambda = 0.693132$. This is, of course, negligibly different from the analytically obtained value of $\bar{\lambda} = 0.693147\ldots \simeq \ln 2$.

[2] The Lebesgue measure is simply the length or volume measure of a system. Given a region $A$, the Lebesgue measure $\mu(A)$ is given by $\mu(A) = \int_A dx$.

Lyapunov exponents can also be easily calculated for (44) for values of $p < 3$. When $p = 2$, there is a stable fixed point at $x = \frac{1}{2}$; furthermore, (44) is a contracting map on the domain $(0, 1)$, since $D_x f > 0$ for $x < \frac{1}{2}$ and $D_x f < 0$ for $x > \frac{1}{2}$. Any two trajectories of the system in the domain $(0, 1)$ therefore converge to the fixed point $x = \frac{1}{2}$, from which we immediately deduce that the Lyapunov exponent $\lambda$ of the system cannot be positive. The same argument can be given for any value of $p < 3$. Note that does not matter whether $\lambda$ is less than zero or simply equal to zero - just the fact that it is non-positive confirms the non-chaotic nature of the map for these cases.

### 4.1.2 Ergodicity

Definition (5) can be used to demonstrate the ergodicity of the logistic equation. Note that the tent map (46) can be reexpressed as the Bernoulli shift:

$$y_{n+1} = \frac{\arcsin^2(x_{n+1})}{\pi} = f(y_n) = 2y_n \ (\text{mod} 1). \tag{49}$$

Consider a piecewise continuous (and continuously differentiable) function $g : \mathbb{R} \to \mathbb{R}$. $g$ has a Fourier expansion:

$$g(y) = \sum_{n=-\infty}^{+\infty} a_n e^{2\pi i n y}. \tag{50}$$

If $g \circ f = g$, as required by definition (5), then we need

$$g \circ f(y) \equiv g(2y \ (\text{mod } 1)) \equiv \sum_{n=-\infty}^{+\infty} a_n e^{2\pi i n y} e^{2\pi i n y} mod(1) = \sum_{n=-\infty}^{+\infty} a_n e^{2\pi i n y}, \tag{51}$$

or $a_n = a_n e^{2\pi i n y}$ for all $n$. However, $e^{2\pi i n y} \neq 1$ unless $n = 0$, and therefore we have $a_n = 0$ for all $n \neq 0$. This further implies $g(y) = a_0$, which satisfies definition (5) and demonstrates that (44) is ergodic.

### 4.1.3 Topological Entropy

Consider an initial partition $A$ of $[0, 1]$ consisting of the single set $a_1 = [0, 1]$. We then have $f^{-1}(A) = B = \{b_1, b_2\}$ where $b_1 = [0, .5]$ and $b_2 = [0.5, 1]$ . Continuing, we see $f^{-2}(A) = f^{-1}(B) = C = \{c_1, c_2, c_3, c_4\}$ as illustrated in figure 4.

From this, we may deduce that $A \vee f^{-1}(A) = \{b_1, b_2\} = A^2$, and that $A \vee f^{-1}(A) \vee f^{-2}(A) = \{c_1, c_2, c_3, c_4\} = A^3$. It is easy to see that this doubling continues indefinitely, yielding:

$$\begin{aligned} \Rightarrow \quad N(A^n) &= 2^{n-1} \\ \Rightarrow \quad h(f, A) &= \lim_{n\to\infty} \frac{1}{n} \ln(2^{n-1}) = \ln 2 \end{aligned} \tag{52}$$

It is also apparent that $A^k$ can be used as a refining sequence $A_k$. Using (34), we obtain

$$\begin{aligned} h(f, A_k) &= \lim_{n\to\infty} \frac{1}{n} \ln(2^{k+n-1}) = \ln 2 \qquad (\forall\, k) \\ \Rightarrow \quad h(f) &= \ln 2. \end{aligned} \tag{53}$$

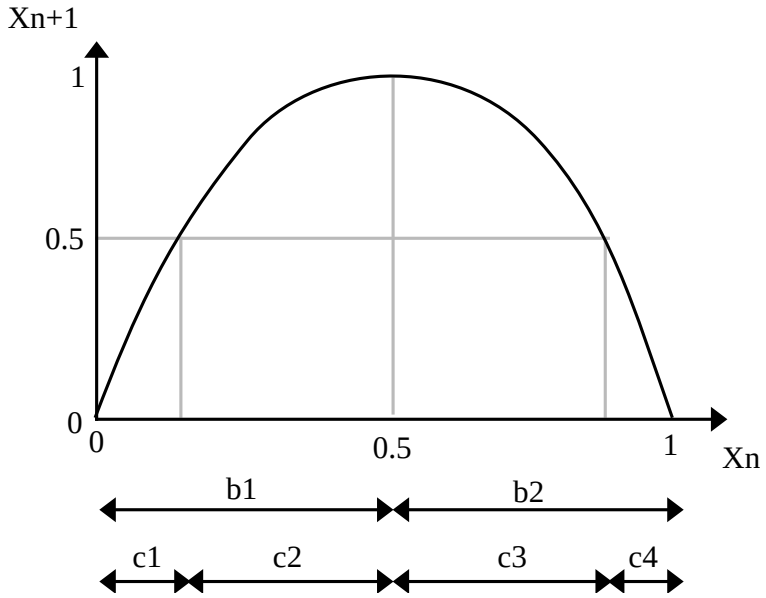



Figure 4: Calculating the topological entropy of the logistic equation

It is interesting to note that the Pesin identity (36) gives the correct value for the measurable entropy, even though it is not applicable as presented due to the fact that the logistic equation is not 1-to-1. We can also observe that the topological entropy calculated above is equal to the measurable entropy; this is not surprising given that $\mu$ is the unique ergodic measure for the $p = 4$ logistic map.

## 4.2 Flows on a Torus

A flow on a torus is a standard model of ergodicity, due both to its simplicity and its applicability to a wide variety of systems. The system given by flows on a torus is Hamiltonian, and can be motivated through the consideration of coupled linear oscillators. As such, its behaviour is relevant to many physical systems as an idealization of coupled oscillators in the uncoupled limit.

Consider a system of coupled linear oscillators given by:

$$\begin{aligned} m\ddot{x} &= k'(y-x) - kx \\ m\ddot{y} &= k'(x-y) - ky. \end{aligned} \tag{54}$$

Using the normal coordinates $a_1 = \frac{y+x}{2}, b_1 = \frac{y-x}{2}$ decouples this system, resulting in the 4 dimensional system of equations (with $w_1^2 = \frac{k}{m}$ and $w_2^2 = \frac{k'}{m}$):

$$\begin{aligned} \dot{a}_1 &= a_2,\ \dot{a}_2 = -\omega_1^2 a_1 \\ \dot{b}_1 &= b_2,\ \dot{b}_2 = -\omega_2^2 b_1. \end{aligned} \tag{55}$$

The further substitution

$$\begin{aligned} a_1 &= \sqrt{\tfrac{2I_1}{\omega_1}} \sin\theta_1,\ a_2 = \sqrt{2\omega_1 I_1}\cos\theta_1 \\ b_1 &= \sqrt{\tfrac{2I_2}{\omega_2}} \sin\theta_2,\ b_2 = \sqrt{2\omega_2 I_2}\cos\theta_2 \end{aligned} \tag{56}$$

yields our final system of equations:

$$\dot{I}_1 = \dot{I}_2 = 0 \tag{57}$$

$$\dot{\theta}_1 = \omega_1,\ \dot{\theta}_2 = \omega_2. \tag{58}$$

The constants $I_1$ and $I_2$ from (57) specify the energies of the two coupled masses, and fixing both restricts the motion of the system to a 2-dimensional integral manifold embedded in the 4-dimensional phase space. (58) tells us that this integral manifold is equivalent to 2-torus, as illustrated by figure 5, and that the system moves on a straight line of slope $\frac{\omega_1}{\omega_2}$ on the unit square that is isomorphic to this 2-torus.

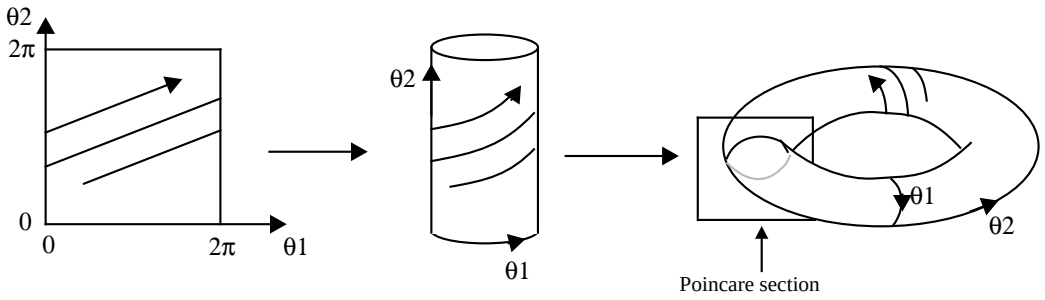



Figure 5: The isometry between the unit square and a 2-torus

### 4.2.1 Lyapunov Exponent

Calculating the Lyapunov exponents of flows on the 2-torus is trivial. Neighbouring flows are parallel straight lines on the unit square isomorphism, which implies that the distance between neighbouring flows is constant. This in turn implies that the Lyapunov exponents $\lambda_1$ and $\lambda_2$ are both zero for the 2-torus.

### 4.2.2 Ergodicity

It is possible to demonstrate the ergodicity of flows on the torus through the application of definition (5) - that is, we could prove that there are no invariant functions $g$ of the flow other than constant functions. This can be done by using a technique identical to that use in the case of the logistic equation. Instead of doing so, we will present a more constructive technique.

In order to prove the ergodicity of flows on the torus, we first need to define some terms. Given $\omega_1, \omega_2 \in \mathbb{R}$, then $\omega_1$ and $\omega_2$ are commensurate if $n\omega_1 + m\omega_2 = 0$ for some $n, m \in \mathcal{Z}$. $\omega_1$ and $\omega_2$ are incommensurate if they are not commensurate.

The following lemma is proven in Peterson[12]:

**Theorem 3 (Ergodicity of Rotations of the Circle)** *Suppose the circle $S'$ is rotated through an angle $\alpha$, and $\alpha$ is incommensurate with $2\pi$. If so, then the sequence $S = \{\theta, \theta + \alpha, \theta + 2\alpha, \ldots, \theta + n\alpha, \ldots \ (mod\, 2\pi)\}, (n \in \mathcal{Z})$ is everywhere dense on the circle.*

A second term that needs to be defined is that of a Poincaré section. Consider a vector field or flow near some (non-trivial) orbit. The flow can be seen to be tranverse to a manifold, which varies along the orbit. The transverse manifold is essentially a surface which is not tangent to the local flow. We call this transverse manifold (of co-dimension 1) a *Poincaré section* of the original flow. Consider then an orbit which pierces the Poincaré section at some point $x_0$. If the flow is nearly periodic in nature, then after some time $\delta t$, the flow will again pierce the Poincaré section at another point $x_1$. The map $g(x_n) = x_{n+1}$ that, given a piercing of the section returns the next piercing of that section, is called a Poincaré map. (Note that not all flows have Poincaré maps.)

Consider a Poincaré section of the torus as illustrated in figure 5, given by the cross-section:

$$\{(\theta_1, \theta_2) | \theta_2 = \bar{\theta}\}. \tag{59}$$

The Poincaré map on this section becomes

$$\theta_1 \mapsto \theta_1 + \beta \cdot 2\pi \pmod{2\pi}, \tag{60}$$

where $\beta = \frac{\omega_1}{\omega_2}$. This is, of course, a rotation on a circle and theorem 3 applies.

If $\omega_1$ and $\omega_2$ are incommensurate (ie. if $\beta$ is irrational), then theorem 3 tells us that the map (60) is dense on the Poincaré section. However, since we did not specify at which angle $\bar{\theta}$ the section was taken, this is true for any angle $\bar{\theta}$. Therefore, if $\omega_1$ and $\omega_2$ are incommensurate, a flow on the torus is everywhere dense and therefore ergodic.

If $\omega_1$ and $\omega_2$ are commensurate, ie. $\omega_1 = \frac{m}{n}\omega_2$, then consider the $n$th iterate of the map (60). This is given by

$$\theta_1 \mapsto \theta_1 + n\beta \cdot 2\pi \; (mod \; 2\pi) = \theta_1 + n\frac{m}{n} \cdot 2\pi \; (mod \; 2\pi) = \theta_1. \tag{61}$$

In this case, all flows on the torus are closed and periodic. This implies that if $\omega_1$ and $\omega_2$ are commensurate, a flow on the torus is not ergodic.

### 4.2.3 Topological Entropy

We have demonstrated that (for irrational slopes $\frac{\omega_1}{\omega_2}$) flows on the torus are ergodic, and we have also calculated the Lyapunov exponents of these flows. We are now in a position to use the Pesin identity (36) and immediately deduce that $h_{meas}(f) = 0$. This does not give us any information regarding the topological entropy of this flow, since measurable entropy only provides a lower bound.

It is, however, possible to calculate topological entropy directly. Consider an initial partition $A_p$ of the unit square into a grid of $p \times p$ small squares with sides of length $\frac{1}{p}$. There are thus $p^2$ regions in this initial partition. Note that this is also a refining partition as $p \to \infty$.

We can then discretize the flow on the torus into steps that are one unit of time apart. This has the effect of simply translating the grid. Then, as illustrated in figure 6, we have $(2p)^2$ regions in $A_p \vee f^{-1}A_p$. Continuing, we find

$$N(A_p \vee f^{-1}A_p \vee \ldots \vee f^{-n}A_p) = (np)^2, \tag{62}$$

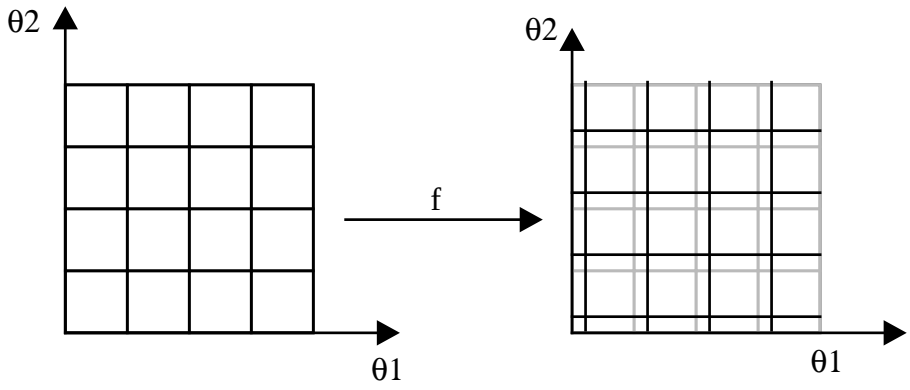



Figure 6: Topological entropy of flows on the 2-torus

assuming that this is the case of irrational slope. We therefore have

$$h(f, A_p) = \lim_{n\to\infty} \frac{1}{n} \ln N(A_p \vee f^{-1}A_p \vee \ldots \vee f^{-n}A_p) = \lim_{n\to\infty} \frac{2}{n} \ln(np) = 0. \tag{63}$$

Thus,

$$h(f) = \lim_{p\to\infty} h(f, A_p) = 0. \tag{64}$$

### 4.3 Two Dimensional Hénon Map

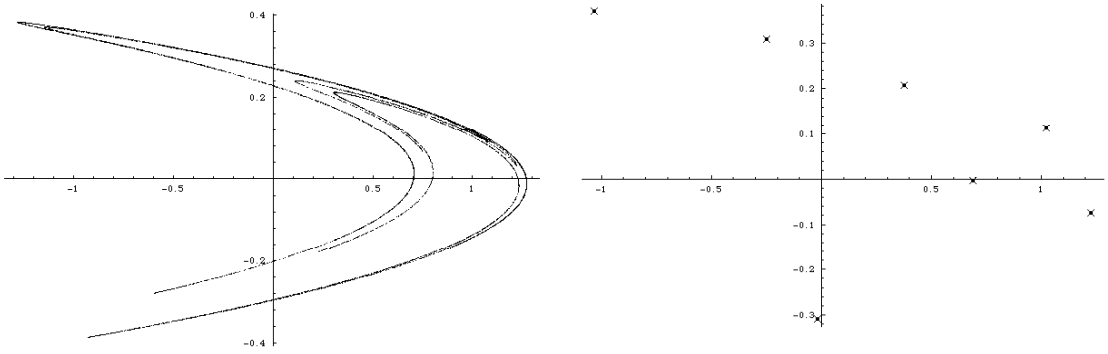
The two-dimensional Hénon map is given by the difference equations:

$$\begin{array}{rcl} x(t+1) & = & y(t) + 1 - a[x(t)]^2 \\ y(t+1) & = & bx(t). \end{array} \tag{65}$$

The behavior of the Hénon map is currently under study. It should be emphasized that this model is not completely understood, but is interesting for precisely this reason. Characteristics such as Lyapunov exponents and entropy have only been approximated numerically - analytical results have not been obtained. However, these numerical studies have demonstrated that this model does exhibit chaoticity for various different values of the parameter $a$. Researchers usually fix the parameter $b$ at a constant value of 0.3, and vary the parameter $a$ in order to generate different behaviours of the map. Consider the Jacobian of (65):

$$J = \begin{bmatrix} -2ax(t) & 1 \\ b & 0 \end{bmatrix} \quad \Rightarrow \quad |J| = -b. \tag{66}$$

(65) thus performs a contraction of area by the factor $b$, while reversing the orientation of the map. For $(a, b) = (1.4, 0.3)$, the system demonstrates chaotic behavior, and the original domain of the mapping is attracted to a fractal set of lines that is reminiscent of a Cantor set (figure 7a). For $(a, b) = (1.3, 0.3)$, the system no longer demonstrates chaoticity, but rather has an attracting orbit of period 7 (figure 7b).



(a) Henon map for (a,b) = (1.4, 0.3)

(b) Henon map for (a,b) = (1.3, 0.3)

Figure 7: The Hénon Map

#### 4.3.1 Lyapunov Exponent, Ergodicity

Ruelle[14] cites a numerical approximation of the Lyapunov exponent $\lambda_1$ for the case $(a, b) = (1.4, 0.3)$ - this result is that $\lambda_1 = 0.42$. We can then use equation (14) and equation (66) to deduce that:

$$\lambda_2 = \ln |J| - \lambda_1 = \ln 0.3 - 0.42 = -1.624. \tag{67}$$

For the case $(a, b) = (1.3, 0.3)$, the periodicity and continuity of the map immediately gives us $\lambda_1 = \lambda_2 = 0$.

The ergodicity of the Hénon map is somewhat inapproachable for the chaotic case. (For the periodic case, we can immediately determine that the map is not ergodic except on the 7 point periodic orbits, which are sets of measure zero.) However, it is known from numerical studies that almost all points are attracted to the Cantor-like set illustrated in figure 7a; because of this, we will optimistically assume that the map is ergodic on this set.

#### 4.3.2 Measurable Entropy

It is not known whether or not we can apply the Pesin identity to this system; in order to apply it, the conditions of the Pesin identity require that we must demonstrate the existence of an absolutely continuous ergodic measure. While an ergodic measure is guaranteed to exist, it may not be related to any physical measure or even the usual Lebesgue measure that one would most naturally associate with the system. However, for the sake of exploring the results, we will simply hope that the identity does apply to the system. Doing so yields

$$h_{meas}(f) = \sum_{\lambda_i > 0} \lambda_i = 0.42 \tag{68}$$

for the case that $(a, b) = (1.4, 0.3)$, and that

$$h_{meas}(f) = \sum_{\lambda_i > 0} \lambda_i = 0.0 \tag{69}$$

for the case of $(a, b) = (1.3, 0.3)$. It is not feasible to analytically calculate the topological entropy of this mapping; the behavior is too complex to be approached. While the measurable entropy can be used as a lower bound on

the topological entropy, a direct numerical method for would be greatly preferred, especially considering the dubiousness of the application of the Pesin identity - such a direct numerical method is discussed in section (5).

## 4.4 Hyperbolic Manifolds

A common problem encountered in general relativity is to find geodesics on a curved manifold. Geodesics are the paths travelled by freely falling particles. Given a unit tangent $u^a = \frac{dx^a}{d\tau}$ with $u^a u_a = 1$, the trajectory of a geodesic is specified by:

$$u^b \nabla_b u^a = 0 \tag{70}$$

or, equivalently,

$$\ddot{u}^a + \Gamma^a_{bc} u^b u^c = 0 \tag{71}$$

where $\Gamma^a_{bc}$ are the Christoffel symbols associated with a given metric.

Given the metric of a manifold, it is possible to determine the geodesic curves by using the well-known fact that geodesics are curves on the space which locally minimize distance. This description of a geodesic lends itself well to the Euler-Lagrange minimization of action integrals. (A discussion of the this theorem can be found in [6].) We can equivalently choose to extremize distance squared, in which case our action integral becomes:

$$I = \int_1^2 g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta d\tau \tag{72}$$

where 1 and 2 are the endpoints of the variational path on the manifold, and $g_{\alpha\beta}$ is the metric of the manifold. The resulting extremal path is a geodesic of the manifold.

One particular set of 2-manifolds (the hyperbolic manifolds) is particularly suitable for study both because of the relative ease with which they can be analyzed, and because the geodesics of these manifolds exhibit chaotic behavior. Hyperbolic spacetimes do appear when solving Einstein's field equations, and also can be found in 2+1 gravity. Hyperbolic manifolds can also be thought of as generic, in some sense. In two dimensions, the integrated curvature of a 2-manifold can be related to its Euler characteristic:

$$\int_{\mathcal{M}} \sqrt{g} d^2 x R = 4\pi\chi. \tag{73}$$

The Euler characteristic of a compact, orientable 2-manifold is $\chi = 2(1-g)$, where $g$ is the genus of the 2-manifold, and therefore the Euler characteristic (and thus the curvature) is negative for all orientable 2-manifolds of genus 2 or more. There is only one orientable 2-manifold ($S^2$) with $\chi = 2$ that shows positive curvature and one orientable 2-manifold ($T^2$) with $\chi = 0$ that has zero curvature. All other 2-manifolds (of which there are a countably infinite number) have strictly negative curvature. Thus, selecting a random 2-manifold from the space of all 2-manifolds will (almost always) result in the selection of a hyperbolic manifold. (See references in Schleich[15].)

Two common and useful models of hyperbolic space are:

1. The hyperbolic half-plane $\{(x, y) \mid y > 0\}$ with the metric

$$ds^2 = \frac{1}{y^2}\left(dx^2 + dy^2\right). \tag{74}$$

Locally, the metric on any closed, hyperbolic manifold is the same as that of the hyperbolic half-plane.

2. The Poincaré disk (also known as Lobachevskian space), given by $\left\{(x,y) \mid x^2+y^2 < \frac{1}{a}\right\}$, with the metric

$$ds^2 = \frac{dx^2+dy^2}{1-a(x^2+y^2)} \tag{75}$$

Upon selecting the metric (74) of the manifold to be our action, the Euler-Lagrange equations yield:

$$\frac{d}{d\tau}\frac{2\dot{x}}{y^2} = 0 \tag{76}$$

$$-2y^{-3}\left(\dot{x}^2+\dot{y}^2\right) - \frac{d}{d\tau}\frac{2\dot{y}}{y^2} = 0 \tag{77}$$

We can then use (76) to deduce

$$\frac{dx}{d\tau} = Cy^2 \tag{78}$$

and then substitute (78) into (77) to generate

$$\frac{dy}{d\tau} = \left[Ey^2 - C^2y^4\right]^{\frac{1}{2}} = \pm y\left[E - C^2y^2\right]^{\frac{1}{2}}. \tag{79}$$

Alternatively, we define $u^a = \dot{x}x^a + \dot{y}y^a$ as our vector tangent to the geodesics, and then normalize this tangent vector with the condition $u^a u_a = 1$. This yields

$$\frac{1}{y^2}\left(\dot{x}^2+\dot{y}^2\right) = 1 \tag{80}$$

which can be substituted into (78), which results in

$$\dot{y}^2 = y^2 - C^2y^4 \tag{81}$$

Equations (78) and (81) describe the tangents to geodesics of the hyperbolic plane. It is simple to show that these geodesics are semicircles and straight lines that meet the $y = 0$ axis of the plane at right angles, as depicted in figure 8. Refer to appendix A for such a proof.

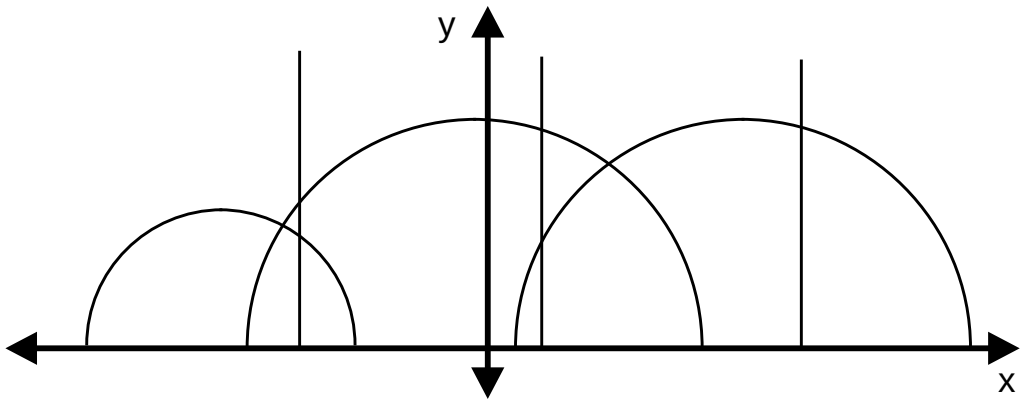



Figure 8: Geodesics of the hyperbolic plane

### 4.4.1 Lyapunov Exponents

In section 2.2 of this document, Lyapunov exponents are described as being a measure of the exponential separation of two trajectories of a given system. For our hyperbolic manifolds, the trajectories are simply geodesics. There is a well-known formulation for calculating the separation of geodesics - namely geodesic deviation - that we can adapt in order to calculate the Lyapunov exponents of our hyperbolic plane geodesics. Schutz[18], or virtually any other introductory text on general relativity, provides a full derivation of geodesic deviation, but we will just state the final result here.

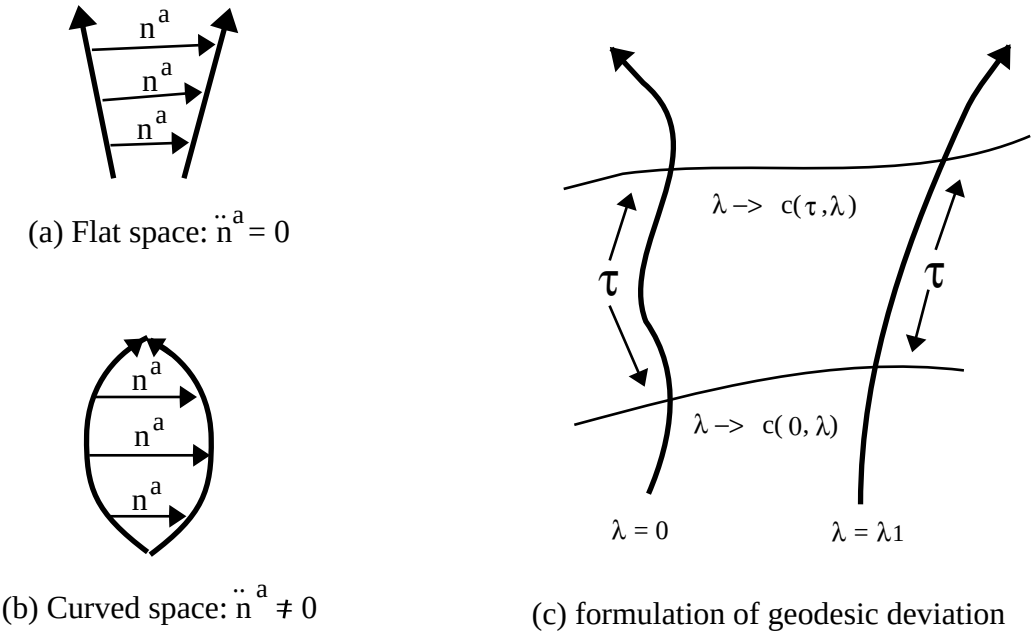



Figure 9: Geodesic deviation

Consider a two-parameter family of curves $c(\tau, \lambda)$, where $\tau \to c(\tau, \lambda)$ are geodesics, and $\lambda \to c(0, \lambda)$ is orthogonal to $\tau \to c(0, \tau)$ at $c(0, 0)$. The curves $\lambda \to c(\tau, \lambda)$ result from dragging the curve $\lambda \to c(0, \lambda)$ a distance $\tau$ up the geodesics. (Refer to figure 9c for an illustration of this.) We refer to the vector tangent to $\lambda \to c(\tau, \lambda)$ as $n^a$: $n^a$ is the connecting vector between the neighbouring geodesics $\tau \to c(\tau, 0)$ and $\tau \to c(\tau, \lambda_1)$. Since $c(0, \lambda)$ is orthogonal to the geodesic, we know that $u^a n_a = 0$.

**Theorem 4 (Geodesic Deviation)** *In this case, the equation of geodesic deviation is given by:*

$$u^b \nabla_b \left(u^c \nabla_c n^a\right) = \ddot{n}^a = \frac{d^2 n^a}{d\tau^2} = R^a_{bcd} u^b u^c n^d. \tag{82}$$

*where $R^a_{bcd}$ is the curvature of the manifold in question.*

In order to compute the curvature we could at this point manipulate (76) and (77) to get:

$$\frac{d^2 y}{d\tau^2} - \frac{2}{y}\frac{dx}{d\tau}\frac{dy}{d\tau} = 0 = \frac{d^2 y}{d\tau^2} + \frac{1}{2y}\frac{dx}{d\tau}\frac{dx}{d\tau} - \frac{3}{2y}\frac{dy}{d\tau}\frac{dy}{d\tau} \tag{83}$$

from which we can extract the Christoffel symbols

$$\Gamma^x_{xy} = \Gamma^x_{yx} = -\frac{1}{y}, \qquad \Gamma^y_{xx} = \frac{1}{2y}, \qquad \Gamma^y_{yy} = \frac{-3}{2y} \tag{84}$$

for the hyperbolic plane. We can use these Christoffel symbols and the definition of curvature

$$R^\alpha_{\beta\mu\nu} \equiv \Gamma^\alpha_{\beta\nu,\mu} - \Gamma^\alpha_{\beta\mu,\nu} + \Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\beta\nu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\beta\mu} \tag{85}$$

to yield the fact that:

$$R_{abcd} = -\kappa\left[g_{ac}g_{bd} - g_{ad}g_{bc}\right] \tag{86}$$

for the hyperbolic plane (where $\kappa > 0$). Equation (82) can now be directly applied:

$$\begin{aligned}
\frac{d^2n^c}{d\tau^2} &= -\kappa\left(g^c_a g_{bd} - g_{ad}g^c_b\right)u^d u^c n^b \\
&= -\kappa\left(u^c u_b - u_d u^d g^c_b\right)n^b \\
&= -\kappa\left(u^c u_b n^b - n^c\right) \\
\frac{d^2n^c}{d\tau^2} &= \kappa n^c
\end{aligned} \tag{87}$$

We can solve the differential equation (87) to discover that:

$$n^a = \left(Ae^{\sqrt{\kappa}\tau} + Be^{-\sqrt{\kappa}\tau}\right)x^a + \left(Ce^{\sqrt{\kappa}\tau} + De^{-\sqrt{\kappa}\tau}\right)y^a \tag{88}$$

for $n^a = n^x x^a + n^y y^a$. We may therefore immediately deduce that a Lyapunov exponent for the hyperbolic plane are:

$$\lambda_1 = +\sqrt{\kappa}. \tag{89}$$

This method of deriving Lyapunov exponents based on the deviation of geodesics on a manifold has potential applications in areas of theoretical physics other than general relativity. From classical mechanics, Jacobi's theorem tells us that solving for the dynamics of a Hamiltonian system is equivalent to finding the geodesics of an appropriate manifold. Stated without proof, Jacobi's theorem (refer to Abraham and Marsden[1]) is:

**Theorem 5 (Jacobi's Theorem)** *The integral curves of a Hamiltonian or Lagrangian with a fixed energy $E$ are the same as the geodesics of the manifold with metric $g_E = (E - V)g$, where $E > V$ and the energy is equal to 1, up to reparametrization.*

It is therefore possible to investigate the sensitive dependence of the Hamiltonian system through the application of the equation of geodesic deviation (82).

### 4.4.2 Ergodicity

Obviously the geodesics of the hyperbolic plane (or the Poincaré disk) are not ergodic. However, it is possible to construct a closed surface from the Poincaré disk that does exhibit ergodicity. Because we have already demonstrated that a hyperbolic manifold will have a positive Lyapunov exponent, such an ergodic surface would be chaotic.

The first stage in the construction of such a surface is to consider the isometry between an octagon and a surface of genus 2. Much in the same way that a square can be twice folded to form a 2-torus (figure 5), an octagon can similarly be folded four times to form a surface of genus 2 (figure 10). We associate opposite sides of the octagon with each other, and note that the eight vertices of the octagon become a single point $P$ when the octagon is folded.

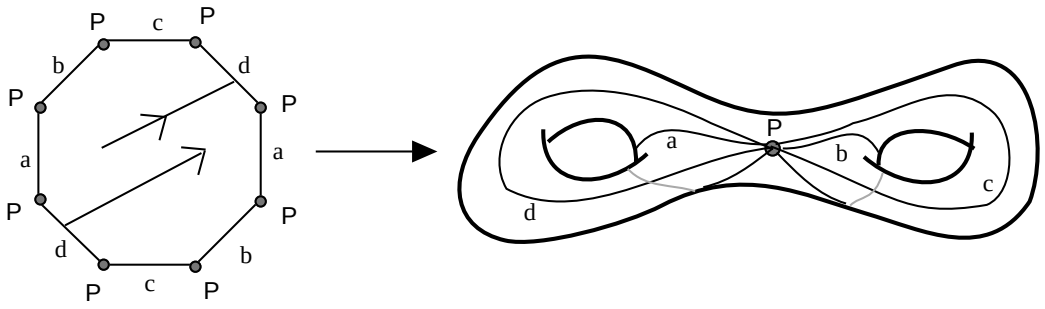



Figure 10: The isometry between an octagon and a surface of genus 2

We now have a suitable topological construction for ergodicity, but by building the octagon out of a region of the Poincaré disk, we can make the construction a smooth manifold with the hyperbolic metric. The next stage in the construction is therefore to select 8 symmetric geodesics around the perimeter of the Poincaré disk. (Refer to figure 11.) As before, each geodesic is circular and intersects the perimeter of the disk with an angle of $\frac{\pi}{2}$ radians. Because our goal is to isolate a region that is topologically equivalent to an octagon, we also need to ensure that $\sum_i \phi_i = 2\pi$. Consider the function

$$\Theta(r) = \sum_i \phi_i. \tag{90}$$

If we let $r \to \infty$, the geodesics become squashed around the perimeter of the Poincaré disk, and the angle $\phi \to 0$. Thus, $\Theta(\infty) = 0$. If we let $r \to 0$, the octagon becomes concentrated about the very center of the Poincaré disk. However, any manifold looks locally like Euclidean space, and in Euclidean space, $\sum_i \phi_i = (n-2)\pi$ for an $n$ sided polygon. Thus, $\Theta(0) = 6\pi$. Because $\Theta(r)$ is a continuous monotonic function, there must be a unique value $r = r', 0 < r' < \infty$ for which $\Theta(r') = 2\pi$, as we require. The resulting surface of genus two has been shown to be ergodic. For a proof of this, refer to Arnold and Avez [3].

We have already seen from equation (73) that it is not possible to construct a 2-dimensional surface of genus 1 (ie. the 2-torus $T^2$) with negative curvature. The 2-torus is isometric to unit square; if we attempt to apply the same technique of constructing the unit square from the Poincaré disk, we should find that the method fails. Applying the method implies we must once again find a radius $r'$ such that $\Theta(r') = 2\pi$. However, since the unit square is a 4 sided polygon, we know that $\sum_i \phi_i = 2\pi$ in Euclidean space, and therefore $\Theta(0) = 2\pi$. Of course, at $r = 0$, there can be no polygon - only a single point. The construction therefore fails as expected for the 2-torus.

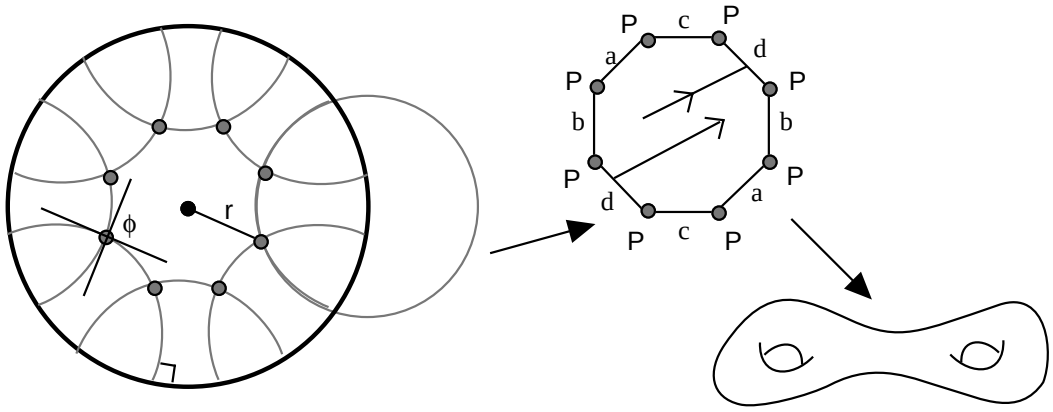



Figure 11: The construction of an ergodic surface with a hyperbolic metric.

# 5 A New Method for the Numerical Approximation of T-entropy

Equation (34) not only serves as a practical tool for calculating topological entropy, but it can also be used to motivate the design of a numerical approximation scheme. If the regions that comprise a partition could be approximated numerically, then it would be possible to evolve these regions (and by extension the partition) through a mapping a large number of times in order to approximate the limit in (34). This approximation can be used as a lower bound for the topological entropy of the mapping; the original partition could then be refined and the process repeated in order to get some idea of the trend observed as the refinement proceeds onward towards infinity.

Such a numerical scheme would be extremely useful as an additional tool in the arsenal for detecting and characterizing chaotic behaviour. The analytical computation of topological entropy is at least as intractable as the calculation of Lyapunov exponents; for this reason alone, a numerical method would be useful. However, no direct method has ever been attempted before because of the computational difficulties posed by the formulation of topological entropy. We will approach the development of such a method by implementing the most naive algorithm, and we will then examine the feasibility and potential pitfalls of a more complex method through the results of this naive attempt.

## 5.1 One-dimensional Maps

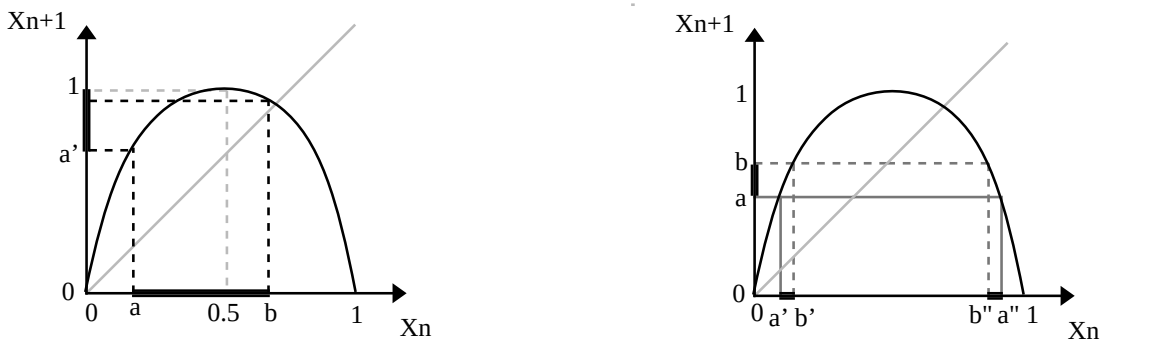
The logistic equation (44) provides an excellent model in order to whet our teeth with a first attempt at numerical approximation. Regions in one dimension are particularly trivial to represent numerically - they are simply intervals, and are specified exactly by two endpoints. The logistic equation also has the useful feature that both chaotic and non-chaotic behaviour can be observed for different values of $p$. The application of a numerical method to these different values of $p$ will indicate whether or not the method is powerful enough to discern between the different behaviours. However, there are a few potential pitfalls that must be mentioned before the algorithm can be specified.

The logistic map is a two-to-one map. Because of this, we must instead use the inverse of the map in order to calculate topological entropy. Given the point $x_{n+1}$, we calculate the inverse of the map as:

$$f^{-1}(x_{n+1}) = x_n = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{x_{n+1}}{p}} \tag{91}$$

Note that this map is (of course) one-to-two. For any given interval $A$, there are two intervals $A_1, A_2$ such that $A_1 = f^{-1}(A)$ and $A_2 = f^{-1}(A)$; both of these intervals must be included in the partition computed by iterating the initial partition's regions through $f^{-1}$.

Note that in addition to dealing with the two-to-one nature of $f$, using $f^{-1}$ also changes the nature of the map in a subtle way. Instead of being expansive in nature, the map is now contractive - the measure of any interval will decrease when mapped through $f^{-1}$ (but there will be two resulting intervals, so the total measure is of course conserved.) This means that the action of $f^{-1}$ itself is refining our initial partition - with any luck we won't have to explicitly consider refinements when choosing the initial partition. Using $f^{-1}$ also means that a potential computational snarl is avoided. If an interval $[a, b]$ containing the point 0.5 is iterated through $f$,then the resulting interval will be $[\min(f(a), f(b)), 1.0]$ and not $[f(a), f(b)]$ as expected. Such a resulting interval will then eventually 'grow' to encompass the entire unit interval, requiring us to pick a finer refinement of the initial partition in order to carry out the approximation any further. Refer to figure 12 for an illustration of these issues.




(a) Taking an interval [a,b] through f

(b) Taking an interval [a,b] through $f^{-1}$

Figure 12: The effect of $f$ and $f^{-1}$ on intervals.

### 5.1.1 Algorithm

We are now in a position to specify the algorithm for estimating topological entropy. The overall approach of the algorithm is to pick an initial partition of $[0, 1]$ into a set of disjoint intervals, and to iterate this initial partition through the inverse logistic equation. At each iteration, the product of the evolved initial partition and a "cumulative product partition" will be computed; this product will then be set as the "cumulative product partition" for the next iteration. The algorithm therefore proceeds through the following steps:

1. **Create the initial partition $IP$ and the initial cumulative product partition $CPP$**

   As we have already discovered, the action of the inverse logistic equation $f^{-1}$ itself will refine any initial partition we choose, so selecting any partition will do the job. For the sake of simplicity, the partition $IP = \{[0, 0.5], [0.5, 1.0]\}$ was chosen in the test implementation.

2. **For each interval in $IP$, evolve it into the next partition $NP$**

   This step is simply the application of equation (91) to each interval in $IP$.

3. **Create $NCPP = CPP \vee NP$**

   The bulk of the computation happens in this step. The implementation of the "$\vee$" operator follows from definition (28). Each interval in $NP$ is intersected with every interval in $NCPP$, and the resulting interval added to $NCPP$ (if this resulting interval is non-empty, of course.) This step requires $n \times m$ operations, were $n$ is the number of intervals in $NP$ and $m$ is the number of intervals in $NCPP$. Note that if the map being used does have positive topological entropy, then the number of intervals in $NCPP$ is expected to grow exponentially. This means that the number of steps required by this operation is expected to be proportional to $\exp(h(f) * n_0 * m_0)$, where $n_0$ and $m_0$ are the initial number of intervals in $IP$ and $CPP$, and $h(f)$ is the topological entropy of the map. We thus expect to see an explosion in the number of computational cycles (and storage space) required for this step after a sufficient number of iterations.

4. **If adequate convergence is obtained, quit, otherwise set $IP = NP$, $CPP = NCPP$, and goto step 2**

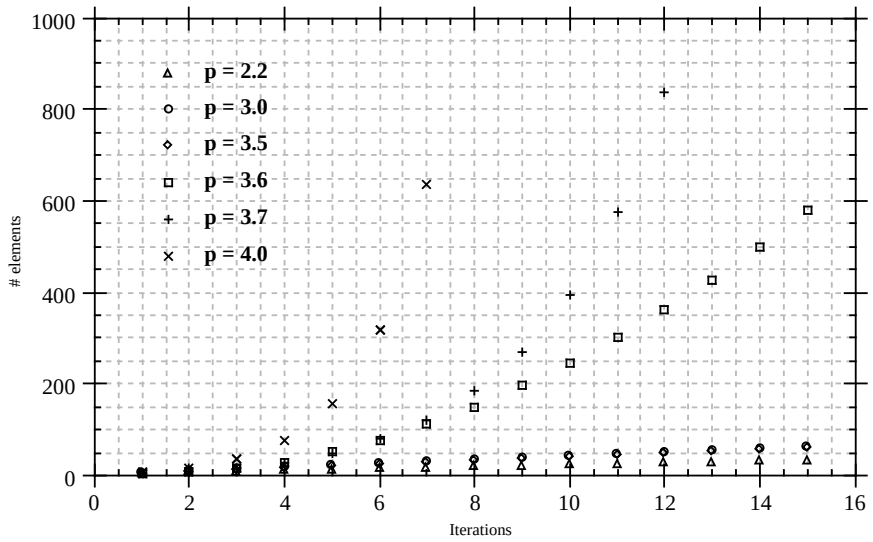
   The concept of "adequate convergence" is difficult to define. Somewhat ironically, we are rescued from having to identify such a definition by the computational explosion that occurs in step 3 after some number of iterations of the algorithm. The best we can do is output the current estimate of $h(f)$ at each step, for as many steps as it takes for the explosion to occur.

   If we were able to push the algorithm through a large number of iterations, a crude test for convergence would be to define $h_i(f)$ as the approximation of topological entropy at the $i^{th}$ iteration, and then require $h_i - h_{i+1} < \epsilon$ for some user-defined $\epsilon$. (Of course, we have not even shown that this algorithm itself converges!)

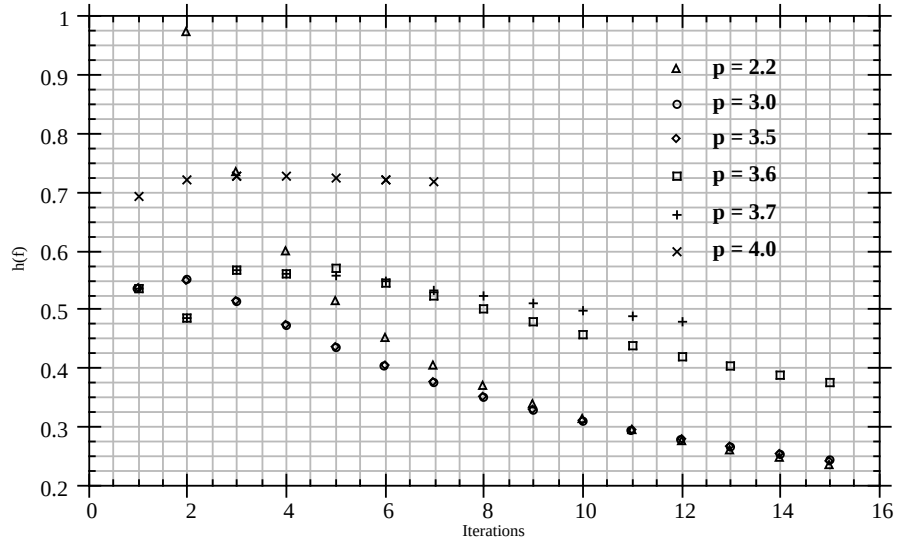
### 5.1.2 Implementation and Results

Table 1 and figure 13 summarize the results of applying an implementation of the approximation algorithm to the logistic equation, for various values of the parameter $p$. Figure 13a illustrates the growth in the number of intervals within $CPP$, while 13b depicts the convergence of the method as the number of iterations increases. It should be noted that it can be analytically demonstrated that the logistic equation does exhibit chaoticity for $p = 4.0, p = 3.7$, and $p = 3.6$, but does not for $p = 3.0$ and $p = 2.2$.

The numerical data strongly supports the analytical predictions, which implies that this simple method is a success. For the $p = 4.0, p = 3.7$, and $p = 3.6$ cases a very strong exponential growth in the number of intervals within $CPP$ can be observed in figure 13a, in contrast to the $p = 3.0$ and $p = 2.2$ cases, in which the growth seems




(a) # elements in P vs. iterations




(b) approximated h(f) vs. iterations

Figure 13: Results of approximating $h(f)$ for the logistic equation

| p | # iterations | # elements in $CPP$ | approximated $h(f)$ |
|---|---|---|---|
| 2.2 | 15 | 33 | 0.23310 |
| 3.0 | 15 | 61 | 0.24182 |
| 3.5 | 15 | 61 | 0.24182 |
| 3.6 | 15 | 580 | 0.37430 |
| 3.7 | 12 | 836 | 0.48062 |
| 4.0 | 7 | 638 | 0.71759 |

Table 1: Numerical results in approximating $h(f)$ for the logistic equation

to be linear or perhaps polynomial at best. Figure 13 is much more illuminating, however. For the $p = 4.0$ and $p = 3.7$ cases, it is clear that the approximate values of $h(f)$ are converging to some constant value. For $p = 2.2$ and $p = 3.0$, it is also seems that the approximate values are converging towards zero, although it is difficult to discern whether or not this is truly headed to zero or just to some relatively small value. The $p = 3.6$ case is more interesting - analytically, we should observe non-zero topological entropy, but numerically it is impossible to determine to what value the series is converging. For this case, more iterations would be needed in order to come to a conclusion, but because of the computational explosion, this is not possible.

A number of useful observations can be made from this one-dimensional approximation. It is clear that one large limitation of this straightforward scheme is the fact that the number of intervals in the partition $f^{-n}(A)$ grows exponentially with $n$ (this is needed for there to be non-zero topological entropy). The amount of memory required to store all of the approximations of these intervals also grows exponentially; available memory could potentially be exhausted after enough iterations. The large amount of data being stored also motivates us to use efficient storage and retrieval structures; since there are no searches being performed in this one-dimensional algorithm, this is not a large issue, but for higher dimensional systems (in which searches become necessary - refer to section 5.2.1 for details) it becomes much more important. We will be discussing such improvements in subsequent sections of this thesis.

In addition to obvious memory resource issues, the exponential growth in the number of intervals in the partitions $IP$ and $CPP$ means that the required number of operations in the algorithm also grows exponentially. It was this computational explosion that limited the implementation to less than 15 steps - after this number of steps, each subsequent iteration would take many hours to perform on a reasonably fast workstation. The success of the algorithm seems to depend on the approximation to $h(f)$ converging to a reasonable value before this explosion takes place; it is still an open question of how to determine the rate of convergence that will take place.

Another issue that needs to be addressed is the fact that although exponential growth is observed, actually demonstrating that the growth is exponential and not polynomial is very difficult. Either a method for doing this needs to be designed, or a more suitable entropy (such as KS-entropy) that has negative as well as zero values needs to be used instead of topological entropy. Having negative values would be a great boon because it is much easier to detect convergence to any value below zero than it is to simply test for convergence to zero itself.

## 5.2 Two-dimensional Maps

Lost in the transition between one and two dimensions is the ability to trivially represent arbitrary regions. Suddenly specifying endpoints is no longer good enough - a more sophisticated representation needs to be found. Computation geometry literature refers to *Voronoi diagrams*, which are very flexible representations that can be generalized to higher dimensions. Sets are represented by vertices $V_i$ - given a vertex $V_j$, the set associated with it is the set of neighbouring points that are closer to the vertex $V_j$ than any other vertex. Unfortunately, the Voronoi representation is not preserved through arbitrary mappings, which makes it unsuitable for our purposes.

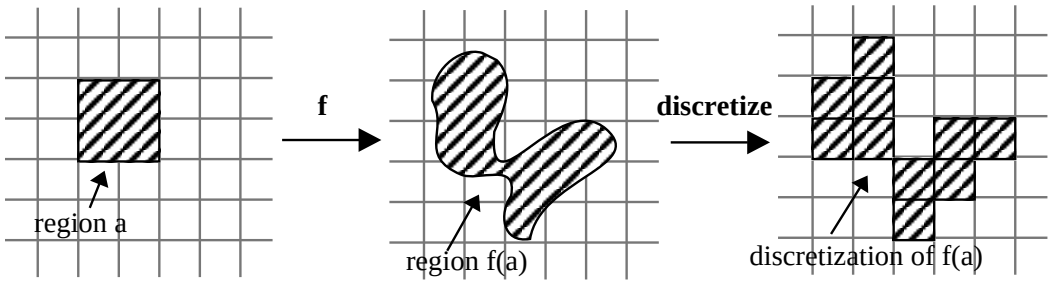



Figure 14: The discretization of regions

A more productive approach is to consider the discretization of regions into a finite number of grid elements (figure 14). Such a discretization is relatively easy to perform, and has the advantage that discretizations of regions can be evolved through a map element by element, and mapped elements can be subsequently discretized to provide an aggregate discretization of the mapping of the region. This approach is not without its disadvantages - inherent in any discretization scheme is the introduction of error. However, this error depends on the resolution of the discretization relative to the size of the region begin approximated, and can therefore be controlled.

Partitions can be considered as a set of regions, and regions are a set of grid elements. We must therefore build a data structure to efficiently store regions, and embed a second data structure to store grid elements inside each member of the first structure. In order to select an appropriate structure, we should first examine what operations will be performed on these structures.

The evolution of a partition $A$ into a new partition $B = f(A)$ is our primary operation. This can be done by evolving each region within $A$, which of course involves evolving each grid element within each region one by one. Because grid elements are rectangles, a coarse approximation of the evolution of the element can be achieved by simply evolving each of the four vertices of the rectangle through the map, resulting in a quadrilateral. This quadrilateral must then be discretized relative to our grid. (Note that because we are expecting the number of elements within a partition to grow exponentially, we need to also increase the resolution of our grid exponentially in order to prevent the grid from "filling up". This introduces rather extreme resource utilization issues into this primitive algorithm.)

Quadrilaterals can be discretized into a finite set of grid elements by clipping the quadrilateral to elements from discretization grid. (The Sutherland-Hodgman clipping algorithm was chosen for our prototype implementation.

Refer to Hill[8], pages 714-718 for a specification of this algorithm.) If the quadrilateral intersects a given grid element, then the result of such a clipping will be a finite set of edges; we can simply test to see if any edges result from clipping the quadrilateral to an arbitrary grid element, and add the grid element to our discretization if the clipping results in a non-zero number of edges. It is also easy to identify the grid elements on which we should perform this clipping test; we can simply identify the four half-planes that bound the quadrilateral, and select all grid elements within this bounding rectangle. Refer to figure 15 for an illustration of this process.

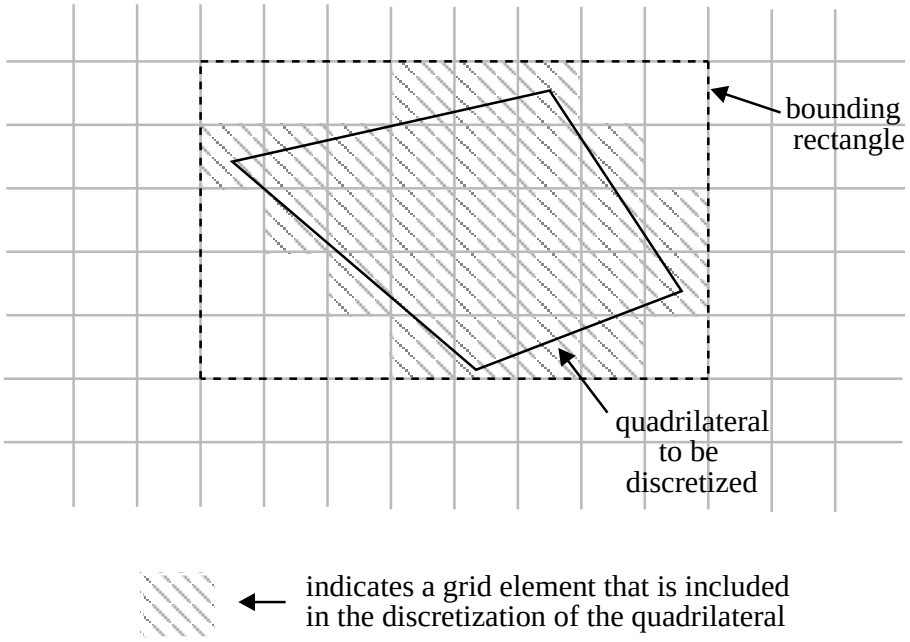



Figure 15: Quadrilateral clipping and discretization

The second operation that needs to be perfomed on two partitions $A$ and $B$ is the computation of their product $C = A \vee B$. We may again appeal to definition 28 of the "$\vee$" operator and reduce the problem to that of computing the intersection of two regions $a_i \in A$ and $b_j \in B$. Since our discretization grid is homogenous, we can simply loop through each grid element in region $a_i$, searching to see if it is also in region $b_j$. If so, then the intersection of $a_i$ and $b_j$ is non-empty, and we add that common grid element to the new region $c_k = a_i \cap b_j$.

As with our one-dimensional algorithm, it is the computation of products that will be the most expensive operation in our two-dimensional algorithm. As well as having to be able to store large amounts of data, we will also have to be able to search for grid elements within regions. This calls for an efficient data structure such as a binary search tree; note that hash tables can not be easily used because of the fact that we don't know how many grid elements will be in any given partition, and thus we don't know how to create a good hashing function or how much space to allocate for the tables. (Refer to appendix C for a description of binary search trees and other relevant data structures.)

It should once again be stressed that this is a rather brute-force algorithm, and is intended more as an

exploration into the issues that will arise from numerical approximations to topological entropy than as an effective method. Hopefully the knowledge gained from this exploration will aid in the design of a more suitable method.

### 5.2.1 Algorithm

Our algorithm for approximating $h(f)$ can now be specified. It is virtually identical to that of section 5.1.1; the major changes occur in the implementation of the evolution and product operations, rather than in the overall algorithm.

1. **Create the initial partition $IP$ and the initial cumulative product partition $CPP$**

2. **For each region in $IP$, evolve it into the next partition $NP$** Regions now need to be evolved and discretized, as described in section 5.2. A suitable polygon clipping algorithm needs to be encorporated into this step; as mentioned before, the Sutherland-Hodgman clipping algorithm was chosen for our prototype implementation.

3. **Create $NCPP = CPP \vee NP$**
   As before, the bulk of the computation happens in this step. The implementation of the "$\vee$" operator is described in section 5.2. For our prototype implementation, height balanced binary-search trees were used for storage and retrieval of both regions and grid elements. (The trees needed to be height-balanced because regions were typically added to the trees in an ordered sequence, which is the worst-case scenario for unbalanced binary search trees. Refer to Cormen, Leiserson, and Rivest [4] for a description of height-balanced binary search tree algorithms, and appendix C for a brief discussion.)

4. **If adequate convergence is obtained, quit, otherwise set $IP = NP$, $CPP = NCPP$, and goto step 2**
   Again, the detection of adequate convergence is academic. The computational explosion associated with the exponential growth in the number of regions with the partition $CPP$ limits the number of iterations of this algorithm that can be performed. The situation is even worse than the one-dimensional case, since we are also required to exponentially increase the resolution of our refinement grid in order to keep up with the growth in the number of regions in $CPP$. This resolution increase results in dramatically more grid elements per region, again causing a computational explosion. We simply don't have the luxury of being able to test for convergence.

### 5.2.2 Implementation and Results

A suitable two-dimensional map that can be used to test this discretization approach is the Hénon map (refer to section 4.3). The Hénon map is dissipative, and contracts an initial region into a lower-dimensional unstable manifold that is similar to a Cantor set in nature. Because of this contraction, and because of the intricate folding of the initial region by the Hénon map, any discretization errors will be rapidly propagated and increased

as more and more iterations through the map are performed. Because of this, the Hénon map will emphasize the limitations of discretization, and will give us a clear idea of what needs to be improved with the approximation scheme. (Note that we anticipate that the difficulties encountered with the Hénon map will be more severe than those encountered with applications in general relativity.)

An initial partition $A$ of the rectangular area $(-1.5, -0.5) \rightarrow (1.5, 0.5)$ into four regions was created. Each of these four regions was discretized to an initial resolution of 0.1 - this means that there were 75 grid elements in each region (figure 16a). Four iterations of our numerical approximation algorithm were then performed, after which time the computation explosion prevented further iterations from taking place. Table 2 summarizes the data obtained for these four iterations, and figure 16 illustrates the evolution and discretization of the initial partition $A$ through the four iterations.

| iteration # | # regions in $CPP$ | approximated $h(f)$ |
|---|---|---|
| 1 | 4 | 1.3863 |
| 2 | 8 | 1.0397 |
| 3 | 32 | 1.1552 |
| 4 | 108 | 1.1705 |

Table 2: Numerical results in approximating $h(f)$ for the Hénon map

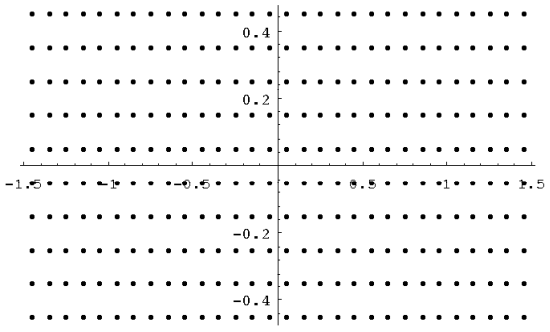
It is fairly evident that this two-dimensional approximation method succumbed to the anticipated difficulties posed by Hénon map. While figure 16 suggests that the numerical scheme preserves the surface characteristics of the Hénon map, the results presented in table 2 indicate that either too much information is being lost in the discretization process, or not enough iterations took place for the approximate value of $h(f)$ to converge.

A closer inspection of the discretization process reveals the fact that the quadrilateral clipping technique is responsible for a large amount of the observed inaccuracy. The method used in the prototype implementation included a grid element in the discretized quadrilateral if any part of the quadrilateral (no matter how small) overlapped the grid element. This has the effect of smearing out the edges of regions, causing neighbouring regions to artificially overlap (figure 17). Not only does this represent a loss of accuracy, it also introduces new, small artifact regions in the product of the partitions $NP \vee CPP$.

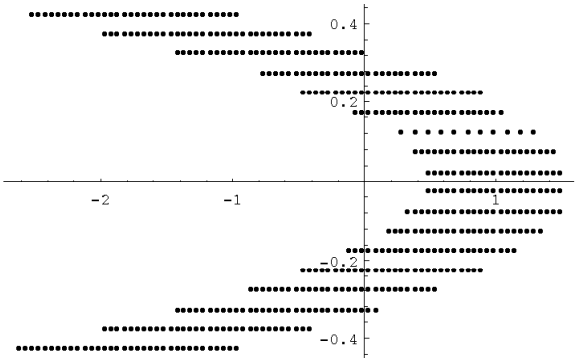
Figure 16 also illustrates the fact that having a globally uniform grid resolution is inefficient. While the edges of regions need to have a fine resolution (in order to properly handle the intersection of regions), the centers of regions could be discretized with a coarser grid. This would greatly reduce the number of grid elements required to discretize a partition, which in turn would allow the algorithm to progress through many more iterations before the computational explosion halts its progress.

## 5.3 Potential Improvements on the Numerical Scheme

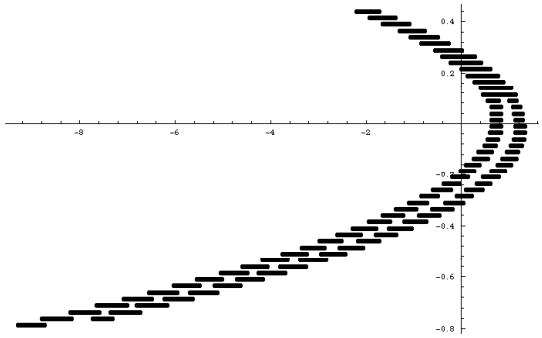
While the initial naive approach to numerically approximating topological entropy seems insufficient, it certainly served its purpose in identifying areas of the algorithm in need of improvement. The inaccuracies involved in the discretization process tend to increase with the number of iterations; this tends to hasten the rate of the



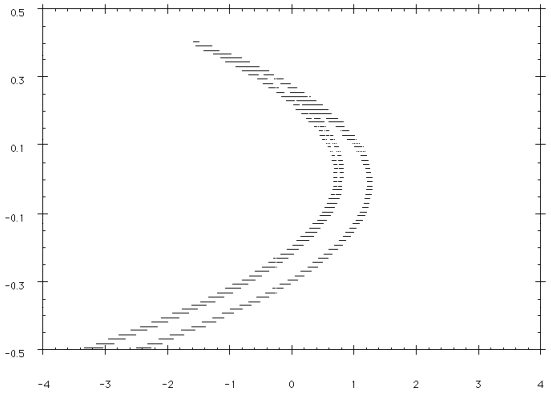

(a) The initial partition A




(b) f(a) - the evolution of A through the Henon map




(c) f(f(a)) - A evolved through the Henon map twice




(d) f(f(f(a))) - A evolved through the Henon map three times

Figure 16: Discretization of the Hénon map

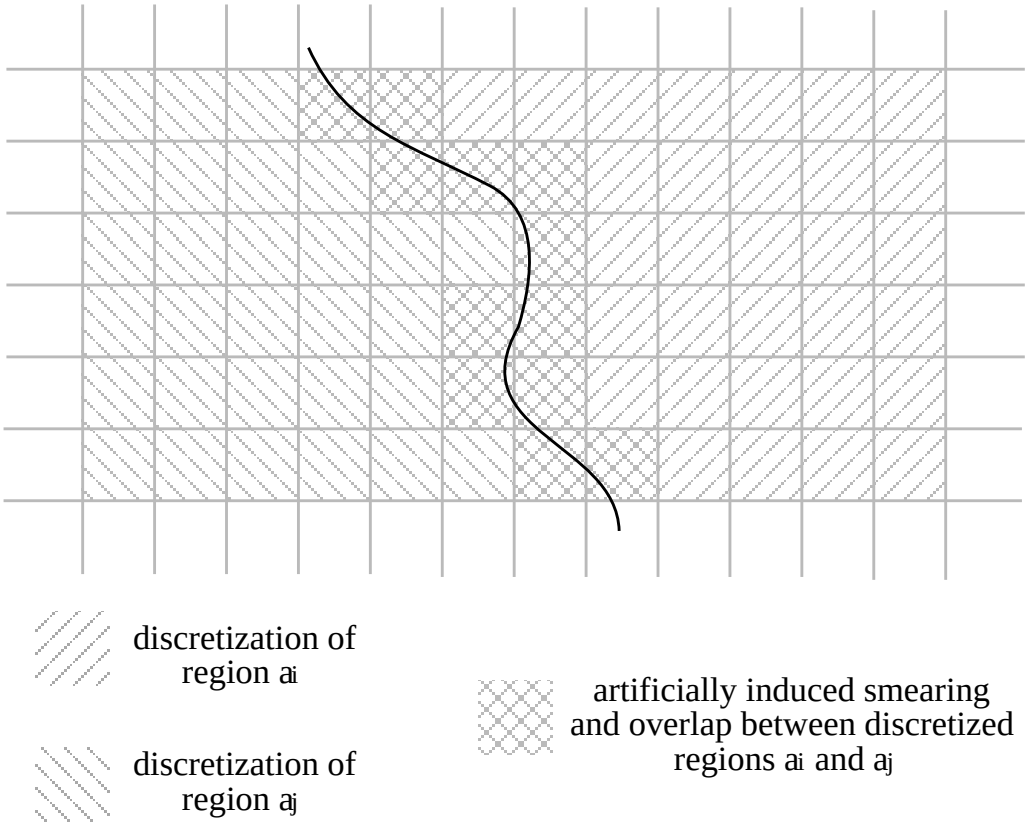



Figure 17: Smearing and overlap of discretized regions

"smearing" of regions' boundaries to the point where the smearing itself dominates the count of regions in the product partition, artificially inflating the approximate topological entropy. The smearing could be controlled by the introduction of a more conservative discretization approach.

Instead of including a grid element in the discretization of a quadrilateral solely on the basis of whether the quadrilateral intersects the element, it would be beneficial to introduce more stringent inclusion requirements. Perhaps the quadrilateral should cover some fraction of the area of the element before the element could be included in the discretization; this would involve the calculation of areas of arbitrary polygons. Such area calculations are possible, and are usually achieved through the triangulation of the arbitrary polygon and then calculating the areas of the individual triangles. This approach would greatly dampen the rate of smearing, increasing the accuracy of the numerical scheme.

The other major difficulty encountered related to the efficiency of the numerical approach, both in terms of storage and computation. Storage became an issue because of the exponential rate of increase in the number of grid elements included in discretizations of regions, which in turn was precipitated by the required increase in resolution. Instead of globally increasing this resolution, it should be possible to use multi-grid techniques to localize the resolution enhancement to the portions of phase space that have the greatest spatial density of regions, and thus require the most representational accuracy.

Both storage and computational requirements could be greatly reduced if the global simulation of the map under study could be appoximated by performing a series of local simulations. Instead of iterating the entire initial partition through the map, the possibility of using a Monte Carlo type approach to limit the scope of the approximation and greatly reduce the amount of computation and storage required.

Consideration also needs to be given to theoretical issues surrounding the proposed numerical scheme. The definition of topological entropy involves the limit $\lim_{n\to\infty}$ of the T-entropy of a the mapping $f$ with respect to an initial partition $A$ (equation 31). As $n$ gets larger, the approximate value of the topological entropy converges towards the final limit; the rate of convergence is obviously a very relevant issue to our numerical scheme. If an analytical estimate of the rate of convergence could be generated, it would be possible to determine how many iterations of the algorithm would be needed in order to gain a given amount of precision in the final approximated value of topological entropy.

# 6 Conclusions and Future Work

In this thesis, the equivalence of the standard Yorke definition of chaos and our definition based on topological entropy has been demonstrated for those systems that don't involve reparameterizations of time. We are not surprised at this result; consider the cases for which the Pesin identity yields $h_{meas} > 0$ (ie. there is at least one positive Lyapunov exponent). Since $h_{meas} \leq h$, it follows that having a positive Lyapunov exponent in these cases implies nonzero topological entropy. Alternatively, if topological entropy $h$ is greater than zero, there exists some measure $\mu$ for which $h_{meas}(\mu) > 0$. Thus, nonzero topological entropy implies that there is at least one positive Lyapunov exponent in such a case.

We have also demonstrated the advantages of using topological entropy as a tool for the characterization of

chaotic systems. Firstly, topological entropy is immune to reparameterizations of time, and as such is an ideal tool for application to general relativity. Furthermore, by using topological entropy, we avoid having to demonstrate the ergodicity of the system under study, which is quite often an exceedingly difficult task. Simply having positive topological entropy is adequate evidence of chaoticity; compare this with the standard definition of chaos, in which the confirmation of a positive Lyapunov exponent is not enough - one must also demonstrate that there is an ergodic measure with support over the entire space. Finally, we have shown that topological entropy is a practical tool by devising a numerical strategy to approximate it.

Although the numerical scheme for the approximation of topological entropy was largely motivated by the inapplicability of the standard definition of chaos to systems studied in general relativity, topological entropy was not calculated (or approximated) for such a system. Hyperbolic manifolds present a perfect opportunity for a first attempt at such a calculation; both an analytical value for the topological entropy of the chaotic hyperbolic manifold presented in section 4.4.2 and a numerical estimate using an algorithm that includes the previously mentioned improvements would be illuminating. Following this, a system actually studied in general relativity (such as the de Sitter spacetime or another such model) should be analyzed using topological entropy.

# A A Demonstration that the Geodesics of the Hyperbolic Half-Plane are Semi-Circles

Equations (78) and (81) can be divided by each other to yield:

$$\frac{dy}{dx} = \frac{\sqrt{y^2 - C^2 y^4}}{Cy^2} = \sqrt{\frac{1}{C^2 y^2} - 1}$$

$$\Rightarrow \frac{dy}{\sqrt{\frac{1}{C^2 y^2} - 1}} = dx \tag{92}$$

We now make the variable substitution $\cosh\theta = \frac{1}{Cy}$. This implies

$$\sinh\theta d\theta = \frac{-dy}{Cy^2} = -C\cosh^2\theta dy.$$

$$\Rightarrow dy = \frac{-\sinh\theta}{C\cosh^2\theta} d\theta \tag{93}$$

We then substitute (93) into (92) to deduce:

$$\int \frac{dy}{\sqrt{\frac{1}{C^2 y^2} - 1}} = -\int \frac{\sinh\theta}{C\cosh^2\theta} d\theta \frac{1}{\sqrt{\sinh^2\theta}} = -\frac{1}{C}\int \frac{d\theta}{\cosh^2\theta}$$

$$= -\frac{1}{C}\tanh\theta = -\frac{1}{C} Cy\sqrt{\frac{1}{C^2 y^2} - 1} \tag{94}$$

We therefore conclude

$$-\frac{1}{C}\sqrt{1 - C^2 y^2} = (x - x_0)$$

or, in a more recognizable form, that:

$$(x - x_0)^2 + y^2 = \frac{1}{C^2} \tag{95}$$

The geodesics of the hyperbolic manifold are therefore semicircles (since we only consider $y \geq 0$); straight line geodesics can be considered to be semicircles of infinte radius, and occur in the case of $C = 0$.

# B Measure Spaces

*The material in this appendix was provided by Schleich andWitt[17].*

It is useful to first define topological spaces before defining measure spaces. Definition: A *topological space* is a set $X$ and a collection of subsets $\mathcal{T}$ satisfying the following properties:

1. $X \in \mathcal{T}$

2. $\emptyset \in \mathcal{T}$

3. If $\{G_\alpha\}$ is a collection of elements of $\mathcal{T}$, then $\cup_\alpha G_\alpha \in \mathcal{T}$ where $\alpha$ is in any index set.

4. If $\{G_k\}$ is a collection of elements of $\mathcal{T}$, then $\cap_k G_k \in \mathcal{T}$ where $k$ is in any *finite* index set.

The elements of the set $X$ are usually called *points*. Further, $\mathcal{T}$ is called a topology on $X$ and the elements of $\mathcal{T}$ are called *open* sets.

Definition: A mapping between topological spaces $f: X \to Y$ is continuous if and only if the inverse image of open sets are open.

Given a continous mapping which is one to one, onto and has a continuous inverse it is called a *homeomorphism*. The two spaces are said to be homeomorphic if there exists sucha mapping and this is an equivalence relation for topological spaces.

Using the above notion of equivalence one can discuss which properties of topological spaces are preserved by homeomorphisms. Such properties are usually called topological. One useful topological property is compactness. A topological space is *compact* if given any collection of open sets which cover it, i.e. every point is in one of the sets, then there are a finite number of open sets in that collection which cover the space. Intuitively, this is means the space is of finite size. An example of such a space is the closed interval 0 to 1, i.e. all real numbers which are greater or equal to 0 and less than or equal to 1. We next define measure space and relate measure spaces to topological spaces through Borel spaces.

Definition: A *measure space* is a set $X$ and a collection of subsets $\mathcal{B}$ satisfying the following properties:

1. $X \in \mathcal{B}$

2. $\emptyset \in \mathcal{B}$

3. If $\{B_k\}$ is a countable collection of elements of $\mathcal{B}$, then $\cup_k B_k \in \mathcal{B}$ where $k$ is in any countable index set.

4. If $\{B_k\}$ is a collection of elements of $\mathcal{B}$, then $\cap_k B_k \in \mathcal{B}$ where $k$ is in any countable index set.

5. If $B \in \mathcal{B}$, then the complement $B^c \in \mathcal{B}$.

The elements of $\mathcal{B}$ are called measurable sets. Note that conditions one, three, and five imply the other two conditions.

In particular, if $(X, \mathcal{T})$ is a topological space and $(X, \mathcal{B})$ a measure space such that $\mathcal{B}$ is the smallest set such that $\mathcal{T} \subset \mathcal{B}$, then $(X, \mathcal{B})$ is a *Borel space* and the measurable sets in $\mathcal{B}$ are called *Borel sets*.

Definition: A map $f : (X_1, \mathcal{B}_1) \to (X_2, \mathcal{B}_2)$ between two measure spaces is called measurable if $f^{-1}(B) \in \mathcal{B}_1$ for all $B \in \mathcal{B}_2$.

One reason for defining Borel spaces is that open sets are all measurable with respect to the Borel measure. Also, any continous map is also measurable with respect the Borel sets.

Definition: A measure $m$ on $(X, \mathcal{B})$ is a function $m : \mathcal{B} \to \mathbf{R}^+ \cup \{\infty\}$ which satisfies the following conditions

1. $m(\emptyset) = 0$

2. Given a countable collection $\{B_k\}$ with $B_i \cap B_j = \emptyset$ for $i \neq j$ implies that

$$m\big(\cup_k B_k\big) = \sum_k m(B_k)$$

where $\mathbf{R}^+$ is the set of non-negative real numbers.

In particular, the measure $m$ is called a *probability measure* if $m(X) = 1$. Further, if the measure is on a Borel space it is called a Borel measure.

One reason for allowing the measure to map some sets to infinity is that measure encountered less abstract situations can have infinite values. For example, the length of the real number line or the volume of a general vector space is infinite.

Definition: Given a measurable map $T : (X, \mathcal{B}) \to (X, \mathcal{B})$. A measure $m$ is a $T$-invariant measure if and only if $T^*m(B) = m(B)$ for all $B \in \mathcal{B}$ where $T^*m(B) = m(T^{-1}(B))$.

Note that $T^{-1}$ is not required to be single valued. Also, it is common for a $T$-invariant measure to be referred to as a invariant measure if it is understood which $T$ is being used.

# C A Brief Discussion of Data Structures

The most basic data structure used for the dynamic storage of a set of nodes is a linked list. A linked list is a data structure in which the nodes of the list are arranged in a linear order. The order in the list is maintained by a pointer in each node, which points towards the next node in the list. In addition to pointing towards the next node in the list, each node also stores some data. Figure 18 illustrates such a linked list of integers. Given a linked list, it is possible to search for a particular node within the list by traversing down the list in a linear fashion. On average, then, it will take $O(N/2)$ traversals to find any given element of the list, if there are $N$ nodes stored in the list. A search operation on a linked list is thus a linear operation.

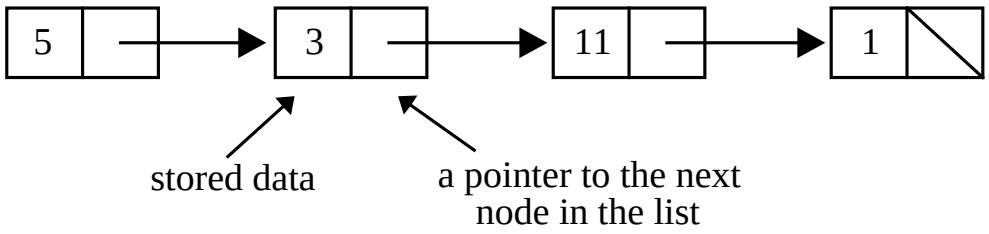



Figure 18: Linked lists

Binary trees are also collections of finite sets of node. A binary tree $T$ either:

- contains no nodes, or
- is comprised of three disjoint sets of nodes: a root node, a binary tree called the *left subtree*, and a binary tree called the *right subtree*.

Each node in a tree contains four fields - data to be stored, a key that is used to order nodes within the tree (this key may in fact be the data), a pointer to the left subtree, and a pointer to the right subtree. The length of the path from the root $r$ of a tree to a node $c$ in the tree is called the depth of $x$ in the tree. The largest depth of any node in the tree is called the height of the tree. Figure 19a provides an illustration of a binary tree.

We can then construct a data structure known as a binary search tree by imposing order conditions on each node. The keys in a binary search tree are always stored in such a way as to satisfy the following property:

> Consider a node $P$ in a binary search tree. If $C$ is a node in the left-subtree of $P$, then $key[C] \leq key[P]$. If $C$ is anode in the right-subtree of $P$, then $key[C] \geq key[P]$. Equality is only needed in these relations if keys are not unique.

Figure 19b depicts a binary search tree.

If there are $N$ nodes in a tree, then on average the height of a randomly built tree will be $O(\ln N)$ nodes. Given a key, it is possible to traverse a binary search tree and find that key by simply tracing a path down from the root of the tree, and comparing the given key with the key of each node on the path. If the given key is greater than the current node's key, then the right subtree is chosen for the next step in the path. If the given

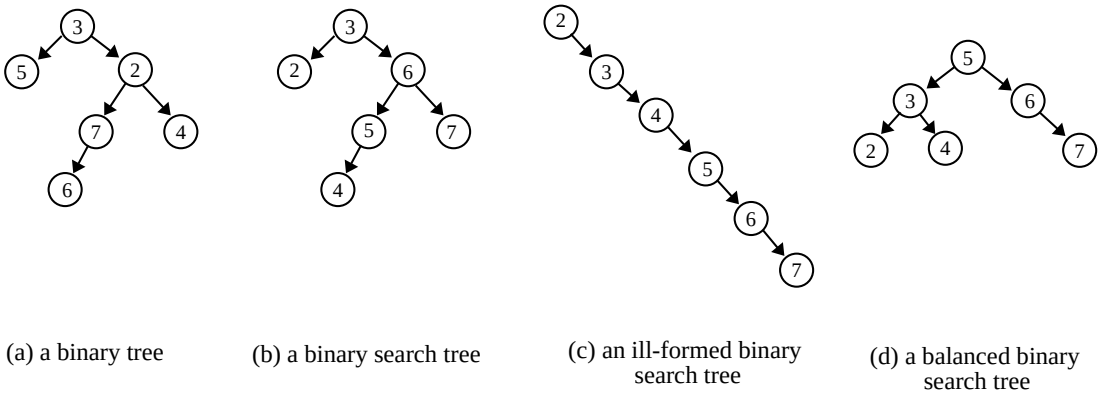



Figure 19: Tree structures

key is less than the current node's key, then the left subtree is chosen for the next step in the path. If the two keys are equal, then the search is successful. With this search strategy, a given key can be found in $O(\ln(N)/2)$ steps. Because of this, binary searching is said to be a logarithmic operation, and is clearly far superior to linear searching for large values of $N$.

It is possible to have an ill-formed binary search tree. Consider the tree depicted in figure 19c; while this tree satisfies our definition of a binary tree, it has the form of a linked list. Searches on this ill-formed tree will be linear, not logarithmic. In order to avoid such a situation, it is possible to define a data structure called a balanced binary search tree, in which given any node $X$ of the tree, the heights of the left and right subtrees of $X$ will differ by at most one. These balanced trees are guaranteed to be well-formed, and will always have the property of logarithmic search time.

# D Source Code - The Lyapunov Exponent of the Logistic Equation

# E Source Code - The Topological Entropy of the Logistic Equation

# F Source Code - The Topological Entropy of the Hénon Map

# References


[1] R. Abraham and J. Marsden. *Foundations of Mechanics, 2nd edition.* Benjamin Cummings, New York, NY, USA, 1985.

[2] R.L. Adler, A.G. Konheim, and M.H. McAndrew. Topological Entropy. *Transactions of the American Math Society,* 114, 1965.

[3] V.I. Arnold and A. Avez. *Théorie ergodique des systémes dynamiques.* 1967.

[4] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms.* The MIT Press, Cambridge, Massachusetts, USA, 1990.

[5] James Gleick. *Chaos: making a new science.* Penguin Books, Harmondsworks, Middlesex, England, 1988.

[6] Herbert Goldstein. *Classical Mechanics (Second Edition).* Addison-Wesley Publishing Company, Reading, Massachusetts, 1980.

[7] Edward Grant, editor. *De proportionibus proportionum by Nicole Oresme.* University of Wisconsin Press, Madison, 1966.

[8] F. S. Hill. *Computer Graphics.* Macmillan Publishing Company, New York, NY, USA, 1990.

[9] E. Atlee Jackson. *Perspectives of nonlinear dynamics, Volume I.* Cambridge University Press, Cambridge, Great Britain, 1991.

[10] E. Atlee Jackson. *Perspectives of nonlinear dynamics, Volume II.* Cambridge University Press, Cambridge, Great Britain, 1991.

[11] Ya. B. Pesin. *Russian Math Survey*, 32:54, 1977.

[12] Karl Peterson. *Ergodic Theory.* Cambridge University Press, Cambridge, Great Britain, 1983.

[13] Mark Pollicott. *Lectures on ergodic theory and Pesin theory on compact manifolds.* Cambridge University Press, Cambridge, Great Britain, 1993.

[14] David Ruelle. *Chaotic evolution and strange attractors.* Cambridge University Press, Cambridge, Great Britain, 1989.

[15] Kristin Schleich. Constraints on the topology of axionic wormhole solutions. *Class. Quantum Grav.*, 9:89–100, 1992.

[16] Kristin Schleich. Personal correspondences, March 1994.

[17] Kristin Schleich and Don Witt. Personal correspondences, April 1994.

[18] Bernard F. Schutz. *A first course in general relativity.* Cambridge University Press, Cambridge, Great Britain, 1990.

[19] Stephen Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer-Verlag, New York, NY, USA, 1990.