BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 44, Number 1, January 2007, Pages 147–155
S 0273-0979(06)01120-7
Article electronically published on October 2, 2006

Computational ergodic theory, by Geon Ho Choe, Springer, Berlin, Heidelberg, New York, 2005, 472 pp., US$89.95, ISBN 3-540-23121-8

When I first read the title of this book I thought it would be about a subfield of ergodic theory, but it is not. It is an introduction to ergodic theory that utilizes computer experiments to illustrate the basic ideas and examples of the subject. Given the historical roots of ergodic theory, this approach seems entirely appropriate. Ergodic theory grew out of problems arising in statistical and Hamiltonian mechanics. In statistical mechanics the motivation is a container filled with a very large number of gas molecules which evolves over time. Since it would be impossible to describe the state of the system using classical differential equations, a probabilistic approach was taken. Now the questions are ones such as, what is the probability that the system is in a given state at a given time? What are the recurrence and mixing properties of the system? What do measurements tell me about the true state of the system? Will the system reach a steady state in the long run?

For a mechanical system with n particles the phase space is a subset of $\mathbb{R}^{6n}$ where each particle is assigned three position coordinates and three momentum coordinates. The state of the system corresponds to a point in the phase space, and the evolution of the system is a trajectory through the space. The evolution of all possible states is a transformation from the phase space to itself. For a Hamiltonian system Liouville's theorem states that the usual $\mathbb{R}^{6n}$ volume is preserved by the transformation. However, in a Hamiltonian system the surfaces of constant total energy are preserved by the transformation so that attention should be restricted to these surfaces. The ergodic hypothesis of Boltzmann was that each surface consists of a single orbit. This is not true, but it was an effort to justify replacing an average of observations along a trajectory over a period of time with the average observation of the entire constant energy surface. One step in such a justification is the ergodic theorem. While each total energy surface does not consist of a single trajectory and not every trajectory visits every part of the surface, there is a measure on the energy surface that is preserved by the transformation. This leads to the study of dynamics of arbitrary flows or transformations on abstract measure spaces.

To begin let (X, μ) be an abstract measure space and assume it is a probability space so $\mu(X) = 1$. Then let T be a flow on X or a transformation of X to itself which preserves the measure. We will assume it is a single transformation and study the discrete time iterates, T^n for $n \in \mathbb{Z}$ when T is invertible and $n \in \mathbb{N}$ when it is not. In this setting measure-preserving means $\mu(T^{-1}(A)) = \mu(A)$ for all measurable sets A. This dynamical system is denoted by (X, μ, T).

The first type of question which can be asked is what are the recurrence properties of T or how does T mix up the space? A basic recurrence property is *ergodicity*. The transformation T is *ergodic* if there are "no" invariant subsets. In terms of measure theory this means that if $T(A) = A$, then the measure of A is zero or one.

2000 *Mathematics Subject Classification.* Primary 37Axx.
In memory of Bill Parry, 1934–2006.

Intuitively it means that T cannot be decomposed into two separate transformations. As an example let (X, μ) be the unit circle with one-dimensional Lebesgue measure and let T be the rotation on the circle by the angle $2\pi\gamma$. If γ is a rational number the transformation is not ergodic, if γ is irrational the transformation is ergodic. *Mixing* is a stronger recurrence property. A transformation is mixing if for any two measurable sets A and B, $\mu(T^n(A) \cap B)$ converges to $\mu(A)$ times $\mu(B)$. It is easy to see that any rotation of the circle fails to be mixing. In P. Halmos' 1956 book [Hal56] he described the difference in the taste of a martini if the stirring is ergodic but not mixing versus the taste if the stirring is mixing. A stronger recurrence property is that of being *independent* or *Bernoulli*. To define this property we define a Bernoulli shift. Suppose $\bar{p} = (p_0, \ldots, p_{k-1})$ is a probability vector. It defines a discrete probability measure on the finite set $\{0, \ldots, k-1\}$ by letting the measure of i be p_i. Define a new probability space to be the product space $\{0, \ldots, k-1\}^{\mathbb{Z}}$ with the product measure $\mu_{\bar{p}}$ from $\bar{p}$. The transformation on the product space is the shift transformation σ defined by $\sigma(x)_i = x_{i+1}$ for all $i \in \mathbb{Z}$. It simply shifts each sequence to the left by one and so preserves the product measure. This measure is independent in that the probability of event i at time zero and event j at time $t \neq 0$ is the product of p_i and p_j. The Bernoulli shift defined by the probability vector $\bar{p}$ is denoted $B(\bar{p})$. The Bernoulli shift $B(1/2, 1/2)$ is thought of as flipping a fair coin. A central idea in ergodic theory is that of isomorphic transformations. Two transformations (X, μ, T) and (Y, ν, S) are *isomorphic* if there is a measure-preserving map $\varphi : X \to Y$ that is a bijection between sets of measure one in each space and $\varphi \circ T = S \circ \varphi$. This means that the transformations are dynamically indistinguishable. A transformation (X, μ, T) is said to be *Bernoulli* if it is isomorphic to a Bernoulli shift. There are intermediate recurrence properties which are also important. *Weak mixing* is a property between ergodicity and mixing, while K is a property between mixing and Bernoulli.

Now we can return to the problem in Hamiltonian dynamics of replacing a time average by a space average. We think of (X, μ, T) as a physical system and an L^1 function $f : X \to \mathbb{R}$ as a measurement or observable. We can make measurements as time evolves, $f(x), f(T(x)), f(T^2(x)), \ldots, f(T^n(x))$ and then ask, what does the average of these measurements tell us about the overall state of the system? That is the *ergodic problem* which concerned Boltzmann. This leads to the ergodic theorem.

Theorem 0.1 (The Ergodic Theorem (G. D. Birkhoff, 1931))**.** *Let (X, μ) be a probability space, $T : X \to X$ a measure-preserving transformation and $f : X \to \mathbb{R}$ an L^1 function. Then*

(1) $\lim_{n\to\infty} (1/n) \Sigma_{k=0}^{n-1} f(T^k(x)) = \bar{f}(x)$ *exists for almost all* $x \in X$*;*
(2) $\bar{f}(T(x)) = \bar{f}(x)$ *almost everywhere;*
(3) $\bar{f} : X \to \mathbb{R}$ *is an* L^1 *function.*

The function $\bar{f}$ in the conclusion of the Ergodic Theorem is invariant under T, $\bar{f}(T(x)) = \bar{f}(x)$ almost everywhere. For each real number r the subset $\{x \in X : \bar{f}(x) \geq r\}$ is invariant for T. Then if (X, μ, T) is ergodic, the function $\bar{f}$ is constant almost everywhere and it is the integral of f over X. The conclusion of the Ergodic Theorem becomes

$$\lim_{n\to\infty} \frac{1}{n} \Sigma_{k=0}^{n-1} f(T^k(x)) = \int_X f d\mu.$$

The left side is the *time average*, and the right is the *space average*. In Hamiltonian mechanics this says that if the transformation restricted to a total energy surface is ergodic, we are assured that the time averages converge to the integral of the function. In many cases the transformation restricted to the total energy surface is ergodic, but in many cases it is not.

In the discussion of recurrence properties there is so far no mention of the speed of mixing, and in the discussion of the Ergodic Theorem there has been no mention of the rate of convergence of the average over time. However, one measure of the randomness of a dynamical system is *entropy*. The following definition of entropy is due to A.N. Kolmogorov [Kol58], [Kol59], and Y. Sinai [Sin59b]. In information theory entropy is thought of as a measure of information content, a viewpoint which will be discussed later. Let (X, μ, T) be a fixed dynamical system and $\mathcal{P} = \{\mathcal{P}_0, \ldots, \mathcal{P}_{k-1}\}$ a finite measurable partition of X. Each $\mathcal{P}_i$ is an atom of the partition. Denote the measure of $\mathcal{P}_i$ by p_i. Define the entropy of the partition to be $H(\mathcal{P}) = -\Sigma p_i \log_2 p_i$. It is a measure of how finely and evenly the partition cuts up the space. For instance, if $\mathcal{P}$ is a partition whose atoms have measures $\{2/3, 1/3\}$, $\mathcal{P}'$ is a partition whose atoms have measures $\{1/2, 1/2\}$ and $\mathcal{P}''$ is a partition whose atoms have measures $\{1/3, 1/3, 1/3\}$, then $H(\mathcal{P}) < H(\mathcal{P}') < H(\mathcal{P}'')$. Next consider the new partition $\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{P})$ which is made up of the intersection of the atoms of $\mathcal{P}$ pulled back n times by T. An atom of the new partition is $\mathcal{P}_{i_0} \cap T^{-1}(\mathcal{P}_{i_1}) \cap \cdots \cap T^{-n+1}(\mathcal{P}_{i_{n-1}})$. The quantity $H(\vee_{j=0}^{n-1} T^{-j}(\mathcal{P}))$ is a measure of how finely and evenly the partition cuts up the space under the action of the transformation. This is computed asymptotically to get the entropy of the partition $\mathcal{P}$ under the action of the transformation T,

$$h(\mathcal{P}, T) = \lim_{n\to\infty} \frac{1}{n} H(\vee_{j=0}^{n-1} T^{-j}(\mathcal{P})).$$

Finally define the entropy of the transformation to be the supremum over all finite partitions

$$h(T) = \sup\{h(\mathcal{P}, T) : \mathcal{P} \text{ finite partition}\}.$$

Two isomorphic dynamical systems clearly have the same entropy. Entropy would be of limited use were it not computable. A partition $\mathcal{P}$ generates the measurable σ-algebra under T if it is the smallest σ-algebra which contains $\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{P})$ for all n. In 1959 Y. Sinai [Sin59b], [Sin59a] proved that if $\mathcal{P}$ is a partition that generates the measurable σ-algebra under T, $h(T) = h(\mathcal{P}, T)$. A basic question at the time was whether or not the Bernoulli shifts $B(1/2, 1/2)$ and $B(1/3, 1/3, 1/3)$ are isomorphic. From Sinai's theorem it follows immediately that if $B(\bar{p})$ is a Bernoulli shift defined by the probability vector $\bar{p} = (p_1, \ldots, p_k)$, then the entropy is $-\Sigma p_i \log_2 p_i$. The entropy of $B(1/2, 1/2)$ is $\log_2 2$ and the entropy of $B(1/3, 1/3, 1/3)$ is $\log_2 3$, which shows they are not isomorphic. Later, in 1970 W. Krieger [Kri70] proved that if (X, μ, T) is an ergodic transformation with entropy $h(T) < \log_2 n$, then there is a partition $\mathcal{P}$ of X with n atoms that generates the measurable σ-algebra under T. The next question became whether or not two Bernoulli shifts with the same entropy are isomorphic. In 1970 D.S. Ornstein [Orn70] proved that two Bernoulli shifts with the same entropy are isomorphic.

A good way to get an intuitive feel for entropy is through the Shannon-McMillan Theorem, which states that for sufficiently large n the measure of most atoms in $\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{P})$ is governed by the entropy. If $\mathcal{P}$ is a finite partition, for sufficiently large n, most $\rho \in \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{P})$ in total measure have $\mu(\rho)$ approximately equal to 2 to the power $-h(\mathcal{P}, T)n$.

Another view of entropy comes from information theory. In 1948 C. Shannon [Sha48] defined the entropy of a Markov source and the capacity of an information channel. Recall the sequence space for a Bernoulli shift, which is $\{0,\ldots,k-1\}^{\mathbb{Z}}$, and the shift transformation on the space. The shift space is a metric space with the metric $d(x,y) = 2^{-\ell(x,y)}$ where $\ell(x,y)$ is 0 if $x_0 \neq y_0$ and is $\max\{\ell : x_i = y_i, \text{for all } |i| < \ell\}$ otherwise. With this topology the shift transformation is a homeomorphism of the space to itself. The sequence space with the metric topology and the shift transformation is called the *full shift on k symbols*. Define a closed, shift invariant subset of the full shift on k symbols by transition rules. The rules forbid certain symbols following others. The subset consists of all sequences that obey the transition rules. For example, start with the sequence space $\{0,1\}^{\mathbb{Z}}$ and forbid the symbol 0 following itself. The space together with the shift transformation is a *topological Markov shift*. Such a shift is conveniently specified by a $k \times k$ zero-one transition matrix. The set just mentioned is specified by $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. The fact that the 00 entry is zero means that the transition is forbidden, and the one in the other ij entries means the other transitions are allowed. Shannon viewed this as a model for a channel through which the allowable strings of symbols can be transmitted. He assumed that the transition matrix is primitive, meaning some power of the matrix is strictly positive. He defined the *capacity* of the channel to be $\lim_{n\to\infty}(1/n)\log_2 N(n)$ where $N(n)$ is the number of allowed sequences of length n. By the Perron-Frobenius theorem this converges to the logarithm of the largest eigenvalue of the transition matrix. This number is now known as the *topological entropy* of the transformation. Next he considered the source of the messages and observed that not all allowable strings might be equally likely, and so he put a Markov probability on the strings of length n to model the statistics of the source. A Markov measure assigns to each transition between symbols a probability. If A is the transition matrix, then a stochastic matrix P of the same size defines such a probability if $P_{ij} > 0$ implies $A_{ij} = 1$. When P is primitive there is a unique probability vector p with $pP = p$. The probability vector defines the probability of the occurrence of a symbol and the matrix P the transition probability. The measure of the set of points with $x_0 = i$ and $x_1 = j$ is $p_i P_{ij}$. Shannon proved that there is a unique number $h(P)$ which satisfies the conclusion of the Shannon-McMillan theorem, and he termed this the *entropy* of the source. He gave the formula $h(P) = -\Sigma p_i P_{ij} \log_2 P_{ij}$, then showed that the entropy of the source is less than or equal to the capacity of the channel, that there is a unique Markov probability whose entropy is equal to the capacity and gave a formula for it. In ergodic theory such a space with the shift transformation and this type of measure is known as a *Markov shift*. The Markov measure that attains the topological entropy is the *measure of maximal entropy*. Independently of Shannon, W. Parry in 1964 [Par64] discovered the measure of maximal entropy and proved it to be unique among all invariant probability measures.

Now consider some examples of measure-preserving transformations.

The first example is known as the *baker's transformation*. The space is $[0,1)^2$, the half-open unit square; the measure is Lebesgue measure; and the transformation is defined by

$$T(x,y) = \begin{cases} (2x, \frac{1}{2}y) & \text{if } x < 1/2 \\ (2x-1, \frac{1}{2}(y+1)) & \text{if } x \geq 1/2 \end{cases} .$$

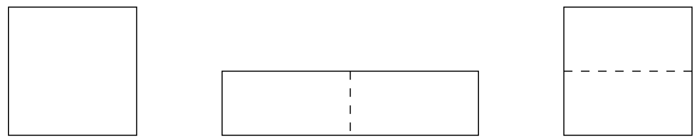

FIGURE 1. Baker's transformation

Pictorially, see Figure 1, the transformation is reminiscent of kneading bread. The entropy of the baker's transformation is $\log_2 2$.

A basic and very important idea is the coding map. If (X, μ, T) is a transformation and $\mathcal{P} = \{\mathcal{P}_0, \ldots, \mathcal{P}_{k-1}\}$ is a partition of X, define the *coding map* $\alpha : X \to \{0, \ldots, k-1\}^{\mathbb{Z}}$ by sending each point to its "itinerary", which means $\alpha(x)_i$ is ℓ if $T^i(x) \in \mathcal{P}_\ell$. Then $\alpha(X)$ is a shift invariant subset of $\{0, \ldots, k-1\}^{\mathbb{Z}}$, it has a shift invariant probability measure $\alpha_*(\mu)$ and $\alpha \circ T = \sigma \circ \alpha$. If the partition $\mathcal{P}$ generates the measurable σ-algebra under T, then the map α is an isomorphism.

For the baker's transformation let $\mathcal{P} = \{\mathcal{P}_0, \mathcal{P}_1\}$ where $\mathcal{P}_0$ is the left half of the square $[0, 1/2) \times [0, 1)$ and $\mathcal{P}_1$ is the right half of the square $[1/2, 1) \times [0, 1)$. The coding map is easily seen to be an isomorphism between the baker's transformation and the Bernoulli shift $B(1/2, 1/2)$.

The next example is similar but noninvertible. The space is the half open interval $[0, 1)$, the measure is Lebesgue measure and the transformation is multiplication by 2 modulo 1. The transformation is measure-preserving because $T^{-1}([a, b))$ is equal to the measure of $[a, b)$ for all intervals. This transformation is ergodic and has entropy $\log_2 2$. Let $\mathcal{P} = \{\mathcal{P}_0, \mathcal{P}_1\}$ be the partition where $\mathcal{P}_0$ is $[0, 1/2)$ and $\mathcal{P}_1$ is $[1/2, 1)$. The coding map $\alpha : [0, 1) \to \{0, 1\}^{\mathbb{Z}_+}$, where Z_+ is the nonnegative integers, sends each point to its dyadic decimal expansion. The ergodic theorem implies that Lebesgue almost every number is normal in its dyadic expansion. The same follows for base 10 using multiplication by 10 modulo 1.

Another family of examples is the toral automorphisms. The n-dimensional torus $\mathbb{T}^n \cong \mathbb{R}^n/\mathbb{Z}^2$ is a compact abelian group where the group operation is inherited from vector addition in $\mathbb{R}^n$. Haar measure on the torus corresponds to Lebesgue measure on the unit cube. An $n \times n$ matrix A with integer entries and determinant plus or minus one acting on $\mathbb{R}^n$ induces a continuous group automorphism T_A of the torus. All continuous automorphisms arise in this way. The matrix preserves Lebesgue measure in the plane, and the automorphism preserves Haar measure on the torus. P.R. Halmos introduced these automorphisms into ergodic theory in 1943 [Hal43] and proved that the automorphism is ergodic if and only if the matrix has no eigenvalues that are roots of unity. The entropy of a toral automorphism is $\log_2 |\lambda_1 \cdots \lambda_t|$, where $\lambda_1, \ldots, \lambda_t$ are the eigenvalues of the matrix A with modulus bigger than one. In 1971 Y. Katznelson [Kat71] proved that every ergodic toral automorphism is isomorphic to a Bernoulli shift. A subclass of the ergodic toral automorphisms are the hyperbolic toral automorphisms which have no eigenvalues of modulus one. Dynamically each point on the torus has a hyperbolic structure under the action of the automorphism. At each point one local subspace contracts exponentially and a complementary subspace expands exponentially. A well-known example on $\mathbb{T}^2$ of such an automorphism is defined by $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and is often referred to as *Arnold's cat map* [AA68] because he used a cat's face in the unit square to

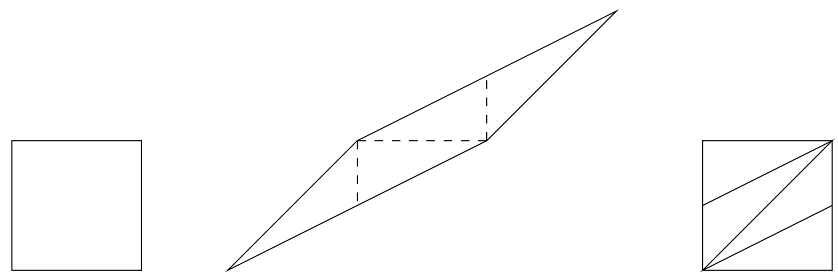

FIGURE 2. Toral automorphism

illustrate the uniform stretching and contracting which occurs under the action of the transformation. It is illustrated in Figure 2.

Hyperbolic toral automorphisms have been important in the study of smooth dynamical systems. Markov partitions were first discovered for automorphisms of the two torus by K. Berg in 1967 [Ber67]and R. Adler and B. Weiss in 1970 [AW70]. They were shown to exist in a much wider variety of cases by Y. Sinai [Sin68] and R. Bowen [Bow70]. A Markov partition is a partition of the space so that the closure of the image of the coding map is a topological Markov shift. There is a geometric characterization of such partitions which allows one to construct them.

Recall that a rotation on the circle is ergodic but not mixing when the angle of rotation is an irrational multiple of 2π. The entropy of any rotation is zero. Fix an irrational $\gamma < 1$ and consider the rotation by $2\pi\gamma$. Define a partition $\mathcal{P} = \{\mathcal{P}_0, \mathcal{P}_1\}$ of the circle where $\mathcal{P}_0 = \{e^{i\theta} : 0 \leq \theta < 2\pi\gamma\}$ and $\mathcal{P}_1 = \{e^{i\theta} : 2\pi\gamma \leq \theta < 1\}$. Let X_γ be the closure of the image of the coding map in $\{0,1\}^{\mathcal{Z}}$. This is a Stürmian minimal set. A zero occurs once in the itinerary of a point for every trip it makes around the circle. A map of a topological space onto itself is minimal if the orbit of every point is dense in the entire space. The original rotation on the circle is also minimal.

The following example is a Denjoy type map of the circle to itself. Begin with the rotation of the circle above. The orbit of $e^{2\pi 0} = 1$ is to be surrounded by a countable collection of open intervals with total measure one that map linearly onto each other. It is done so that the orbits of all other points are unchanged, form a Cantor set and have measure zero. At the point $e^{2\pi 0} = 1$ insert the interval $\{e^{2\pi\theta} : -1/6 < \theta < 1/6\}$ pushing back the rest of the circle to have measure $2/3$. Then at the points that were $e^{\pm 2\pi\gamma}$ insert open intervals of length $1/4$ centered at the points and push back the rest of the circle to have measure $1/6$. Working along the orbit of $e^{2\pi 0}$ inserting two intervals of length $(1/4)^k$ at the k^{th} stage results in a circle with all of the original points still there but a new countable collection of intervals with total length one inserted. The new map on the circle maps the old points to themselves as in the rotation by $2\pi\gamma$, and the new intervals are mapped to themselves center to center and linearly from there out. The result is a homeomorphism of the circle to itself. The intervals are *wandering*, which means that no image of one of the intervals ever intersects itself. The wandering intervals make up an open dense set of measure one. The rest of the points form a Cantor set of measure zero. The example has zero entropy. The map that collapses each of the intervals to a point maps the circle to the circle, and the resulting map is the original rotation by $2\pi\gamma$.

The final example is a *substitution minimal set.* First define a one-sided sequence of 0's and 1's. Begin with a 0. Then form a new string of 0's and 1's by replacing each occurrence of 0 by 01 and each occurrence of 1 by 10. There will be a limiting one-sided infinite sequence. It is

$$.01101001100101101001011001101001\ldots.$$

This sequence has many names, some of which are the Thue sequence, the Morse sequence, the Thue-Morse sequence and others. Let $\mathcal{W}$ denote the collection of all finite substrings (or words) that occur in the sequence and define a closed, shift invariant subset of $\{0,1\}^{\mathbb{Z}}$. A two-sided sequence is in the set if and only if every word that occurs in the sequence is in $\mathcal{W}$. This set together with the shift transformation is often referred to as the Thue-Morse minimal set. The number of words of given length in $\mathcal{W}$ grows linearly as a function of the length so the entropy of the transformation is zero.

Another problem in ergodic theory is the description of the invariant probability measures for a fixed transformation. The space is often a topological space or a manifold, and the transformation is a homeomorphism or a diffeomorphism. We consider this question for the previous examples. The *support of a measure* on a topological space is the smallest closed set that has measure one

A nontrivial full shift or a (primitive) topological Markov shift clearly has uncountably many invariant measures. It is easy to write down measures whose support is the entire space and measures whose support is a proper subset of the space. However, each has a unique measure with maximal entropy, and its support is the entire space. In fact, we will see that in some sense the full shifts contain "all" invariant measures.

The baker's transformation also has uncountably many invariant measures because any Bernoulli measure on $\{0,1\}^{\mathbb{Z}}$ can be transferred to the unit square using the coding map. The baker's transformation has one measure with maximal entropy and it is Lebesgue measure. Hyperbolic toral automorphisms likewise have uncountably many invariant measures because a Markov partition can be used to transfer measures from a topological Markov shift. Lebesgue measure is also the unique measure with maximal entropy for a hyperbolic toral automorphism.

An irrational rotation of the circle is minimal, preserves Lebesgue measure and has no other invariant measures. A transformation with only one invariant measure is *uniquely ergodic*. A Stürmian minimal set is also minimal and uniquely ergodic.

The Denjoy transformation of the circle described above is something else. It is not minimal because it has wandering intervals, but it is uniquely ergodic. Any invariant measure must give the invariant Cantor set measure one because its complement is a collection of wandering intervals. The map that collapses the intervals onto the irrational rotation would push any other invariant measure down to an invariant measure for the irrational rotation and it would not be Lebesgue measure.

The Thue-Morse minimal set arrived at by using a substitution rule is minimal and uniquely ergodic.

To close the discussion of invariant measures we state the *Jewitt-Krieger Theorem* [Jew70], [Kri72]. It states that every ergodic transformation and its invariant measure live in a full shift. Precisely, any ergodic transformation on any space is isomorphic to a closed, shift-invariant subset of a full shift that is minimal and uniquely ergodic.

In the last thirty years there have been spectacular developments in ergodic theory, and many unexpected connections with other fields have been discovered. There has been the development of Bernoulli theory and orbit equivalence, the theory of finitary equivalence; there has been a tremendous amount of work on the actions of the classical matrix groups and the relationships to number theory; there have been applications to smooth dynamical systems and differential equations concerning hyperbolic and partially hyperbolic systems as well as the theory of Lyapunov exponents and invariant measures. There have been uses in complex dynamics, the study of maps of the interval and renormalization. There has also been the application of recurrence theorems to problems in combinatorial number theory. On the practical side ergodic theory has had an impact on information theory problems, particularly in channel coding and data compression.

The book under review is, as stated at the beginning of the review, an introduction to ergodic theory that utilizes computer experiments to illustrate the basic ideas and examples of ergodic theory. The book is designed so that the Maple experiments are an integral part of the book. The author includes actual Maple programs. There are numerical simulations, symbolic computations and graphics programs. The subjects covered in the book are chosen so that the Maple experiments can give insight into the examples and theorems. Most of the basic ideas and examples of ergodic theory are covered and all have Maple programs to illustrate them. Some more advanced topics are also covered, including homeomorphisms of the circle, Lyapunov exponents, Hausdorff dimension and data compression. It provides a mathematical introduction to ergodic theory coupled with a hands-on experimental approach. It is a well thought-out book and illustrates very well how computer experiments can shed light on many aspects of ergodic theory.

References

[AA68] V.I. Arnold and A. Avez, *Ergodic problems of classical mechanics*, Benjamin, 1968. MR0232910 (38:1233)

[AW70] R. Adler and B. Weiss, *Similarity of automorphisms of the torus*, Memoirs of the American Mathematical Society **98** (1970). MR0257315 (41:1966)

[Ber67] K. Berg, *On the conjugacy problem for k-systems*, Ph.D Thesis, University of Minnesota (1967).

[Bow70] R. Bowen, *Markov partitions for axiom* A *diffeomorphisms*, American Journal of Mathematics **92** (1970), 725–747. MR0277003 (43:2740)

[Hal43] P.R. Halmos, *On automorphisms of compact groups*, Bulletin of the American Mathematical Society **49** (1943), 619–624. MR0008647 (5:40c)

[Hal56] ———, *Ergodic theory*, The Mathematical Society of Japan, 1956. MR0097489 (20:3958)

[Jew70] R. Jewitt, *The prevelance of uniquely ergodic systems*, Journal of Mathematical Mechanics **19** (1970), 717–729.

[Kat71] Y. Katznelson, *Ergodic automorphisms of* $\mathbb{T}^n$ *are Bernoulli shifts*, Israel Journal of Mathematics **10** (1971), 186–195. MR0294602 (45:3672)

[Kol58] A. Kolmogorov, *A new metric invariant for transient dynamical systems and automorphisms in Lebesgue spaces*, Academiia Nauk SSSR, Doklady **119** (1958), 861–864. MR0103254 (21:2035a)

[Kol59] ———, *Entropy per unit time as a metric invariant of automorphisms*, Academiia Nauk SSSR, Doklady **124** (1959), 754–755. MR0103255 (21:2035b)

[Kri70] W. Krieger, *On entropy and generators of measure-preserving transformations*, Transactions of the American Mathematical Society **149** (1970), 453–464. MR0259068 (41:3710)

[Kri72] ———, *On unique ergodicity*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume II, Probability Theory, University of California Press, 1972, pp. 327–346. MR0393402 (52:14212)

[Orn70] D.S. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Mathematics **4** (1970), 337–352. MR0257322 (41:1973)

[Par64] W. Parry, *Intrinsic Markov chains*, Transactions of the American Mathematical Society **112** (1964), 55–66. MR0161372 (28:4579)

[Sha48] C. Shannon, *A mathematical theory of communication*, Bell Systems Technical Journal **27** (1948), 379–473, 623–656. MR0026286 (10:133e)

[Sin59a] Y.G. Sinai, *Flows with finite entropy*, Academiia Nauk SSSR, Doklady **125** (1959), 1200–1202. MR0103257 (21:2036b)

[Sin59b] ———, *The notion of entropy of a dynamical system*, Academiia Nauk SSSR, Doklady **125** (1959), 768–771.

[Sin68] ———, *Construction of Markov partitionings*, Funkcional'nyi Analiz i Ego Prilozheniya zen **2** (1968), 245–253. MR0250352 (40:3591)

Bruce Kitchens
Indiana U. Purdue U. Indianapolis
E-mail address: bkitchens@math.iupui.edu