*ENTROPY, A COMPLETE METRIC INVARIANT FOR AUTOMORPHISMS OF THE TORUS**

By R. L. Adler and B. Weiss

THOMAS J. WATSON RESEARCH CENTER, YORKTOWN HEIGHTS, NEW YORK

Communicated by Mark Kac, April 24, 1967

1. *Introduction.*—A central problem in ergodic theory is the question of metric conjugacy: two measure-preserving transformations φ and φ' of a Lebesgue space are said to be metrically conjugate if there exists a third one θ such that $\theta\varphi\theta^{-1} = \varphi'$. In recent years research on this problem has pivoted about a metric conjugacy invariant known as entropy, an idea adapted from information theory, which assigns to a transformation φ a number $h(\varphi) \in [0, +\infty]$. While it is well known that entropy is not a complete invariant, there has been some evidence for its completeness in the class of Kolmogorov transformations.[1] The purpose of this work is to provide more evidence by showing that entropy is a complete invariant for a certain subclass of Kolmogorov transformations, namely the continuous ergodic automorphisms of the torus.

Main Theorem. *If φ and φ' are continuous ergodic automorphisms of the two-dimensional torus such that $h(\varphi) = h(\varphi')$ then there exists a measure preserving transformation θ such that $\theta\varphi\theta^{-1} = \varphi'$.*

The conjugacy will be constructed by coding between two symbolic dynamical systems, and the proof of the main theorem will be executed by representing automorphisms of the torus as such systems.

2. *Automorphisms of the Torus.*—A continuous automorphism φ of the n-dimensional torus $X = E^n/Z^n$ is associated with a unimodular matrix $\Phi = (\phi_{ij})$, i.e., ϕ_{ij} are integers and det $\Phi = \pm 1$. This transformation preserves Haar measure; and it is ergodic if and only if no characteristic value of Φ is a root of unity. For these transformations there are three kinds of conjugacy: algebraic, $\Theta\Phi\Theta^{-1} = \Phi'$ where Θ is also unimodular; topological, $\theta\varphi\theta^{-1} = \varphi'$ where θ is a homeomorphism; and metric, as described above. It is known[2] that for automorphisms of the torus, topological conjugacy is equivalent to algebraic conjugacy and hence implies metric conjugacy. The entropy $h(\varphi) = \log|\lambda_1\ldots\lambda_k|$ where $\lambda_1,\ldots,\lambda_k$ are all the characteristic values of Φ whose moduli are greater than one.[3] Even in dimension two there are algebraically nonconjugate automorphisms with the same entropy such as the pair $\begin{pmatrix}5 & 2\\ 2 & 1\end{pmatrix}$ and $\begin{pmatrix}5 & 4\\ 1 & 1\end{pmatrix}$. The main theorem shows that such pairs are metrically conjugate, which answers a question raised by Adler and Palais.[2]

3. *Symbolic Dynamical Systems.*[4]—Consider an alphabet $\mathfrak{A} = \{1,\ldots,N\}$ of states and a transition rule given by a matrix $T = (t_{ij})$ of zeroes and ones. The space $\Xi = \Xi(T)$ of symbol sequences $\xi = (\ldots, \xi_{-1}, \xi_0, \xi_1, \ldots)$, $\xi_n \in \mathfrak{A}$, consists of those ξ which satisfy $t_{\xi_n, \xi_{n+1}} = 1$ for all n. Let σ denote the shift on Ξ, i.e., $(\sigma\xi)_n = \xi_{n+1}$; let α be the partition of Ξ into sets $\{\xi \in \Xi | \xi_0 = i\}$, $i = 1, \ldots, N$; and let α^n denote the common refinement of $\alpha, \sigma\alpha, \ldots, \sigma^{n-1}\alpha$. The measurable subsets of Ξ are generated by $\cup_{-\infty}^{\infty} \sigma^n\alpha$, and if μ is a σ-invariant measure defined by a matrix $P = (p_{ij})$ of transition probabilities and a row vector $\pi = (\pi_1, \ldots, \pi_N)$ of

stationary probabilities, then the entropy $h_\mu(\sigma)$ of the shift with respect to this measure is $h_\mu(\sigma) = -\sum \pi_i p_{ij} \log p_{ij}$.

THEOREM. *Let T be irreducible and λ the largest positive characteristic value of T with x and y column and row characteristic vectors associated with λ normalized so that* $\sum x_i y_i = 1$. *If* $p_{ij} = x_j t_{ij}/\lambda x_i$ and $\pi_i = x_i y_i$, *then a shift invariant measure μ is defined and* $h_\mu(\sigma) = \log \lambda$. *The maximal entropy for σ on $\Xi(T)$ attained by any normalized invariant measure is* $\log \lambda$ *and μ is the only such measure yielding this value.*

Proof: The first part of this theorem is an application of the Perron-Frobenius theory.[5] For the second part assume that ν is some other normalized invariant measure for which $h_\nu(\sigma) \geq \log \lambda$. The measure ν can be assumed to be singular with respect to μ; consequently there exists sets E_n, $n = 1, 2, \ldots$ such that E_n is a union of elements of α^n and $\mu(E_n)$, $1 - \nu(E_n) \rightarrow 0$, $n \rightarrow \infty$. For any set E which is a union of some elements of α^n denote by $\#(E)$ the number of these elements; whereupon $\kappa\#(E)/\lambda^n \leq \mu(E)$ where $\kappa = \min \lambda\, x_i y_j$. By assumption $H_\nu(\alpha^n)/n$ decreases to $h_\nu(\sigma) \geq \log \lambda$; therefore $\log \lambda^n \leq H_\nu(\alpha^n) \leq \nu(E_n) \log \#(E_n)/\nu(E_n) + (1 - \nu(E_n)) \log \#(E_n{}^c)/(1 - \nu(E_n)) \leq \nu\,(E_n) \log \mu(E_n)\lambda^n/\kappa\nu(E_n) + (1 - \nu(E_n)) \log (1 - \mu(E_n))\lambda^n/\kappa(1 - \nu(E_n))$. Hence $0 \leq \nu(E_n) \log \mu(E_n) + O(1) \rightarrow -\infty$, $n \rightarrow \infty$, a contradiction.

Associated with a measure-preserving transformation φ on a Lebesgue space (X,m) and a partition γ of X into measurable sets $C_1, \ldots, C_N$ a transition matrix $T = (t_{ij})$ can be defined by $t_{ij} = 1$ if $m(\varphi C_i \cap C_j) > 0$ and $t_{ij} = 0$, otherwise. Let χ: $x \rightarrow \mathfrak{A}$ be defined by $\chi(x) = i$ for $x \in C_i$ and τ: $X \rightarrow \Xi(T)$ by $(\tau x)_n = \chi(\varphi^n x)$.

COROLLARY. *If the largest positive characteristic value of T is $e^{h_m(\varphi)}$ and γ is a generator, i.e.,* $\bigcup_{-\infty}^{\infty} \varphi^n\gamma$ *generates the measurable subsets of X, then τ effects a conjugacy between (X,m,φ) and $(\Xi(T),\mu,\sigma)$ where μ is the measure of maximal entropy.*

Proof: The mapping τ is one-to-one a.e. because γ is a generator. By definition $\sigma\tau = \tau\varphi$. Since $h_{\tau m}(\sigma) = h_m(\varphi)$, the theorem implies $\tau m = \mu$.

4. *Coding.*—Consider the space $\Xi = \Xi(T)$ consisting of sequences of symbols from an alphabet labeled by $\mathfrak{A} = \{a_1{}^1, \ldots, a_p^1, a_1{}^2, \ldots, a_q{}^2, b_1{}^1, \ldots, b_r{}^1, b_1{}^2, \ldots, b_s{}^2\}$ whose T is defined by: $t_{ij} = 1$ for $1 \leq i \leq p + q$, $1 \leq j \leq p$, or $p + q < j \leq p + q + r$; $t_{ij} = 1$ for $p + q < i \leq p + q + r + s$, $p < j \leq p + q$, or $p + q + r < j \leq p + q + r + s$; $t_{ij} = 0$, otherwise. The characteristic equation of T is $y^{p+q+r+s-2}\,(y^2 - (p + s)\,y + (ps - qr)) = 0$. Consider also the space $\Xi' = \Xi(T')$ of sequences of symbols from an alphabet $\mathfrak{A}' = \{1,\ldots,N\}$ where T' is defined by: $t'_{1,N} = 0$; $t'_{ij} = 1$, otherwise. The characteristic equation of T' is $y^{N-2}(y^2 - Ny + 1) = 0$.

THEOREM. *If* $ps - qr = 1$ *and* $p + s = N$, *then there exists θ which effects a conjugacy between (Ξ,μ,σ) and (Ξ',μ',σ') where μ and μ' are measures of maximal entropy.*

Proof: After relabeling $\mathfrak{A}'$ by $\{A_1,\ldots,A_p, B_1,\ldots,B_s\}$, the first step in constructing θ is to set $(\theta\xi)_n$ equal to the upper-case letter corresponding to the lower-case one for ξ_n. In order to assign proper indices on the letters, all blocks of b in ξ which are preceded and followed by a are grouped together along with the following a; e.g., $\xi = (\ldots,a,[b,b,\ldots,b,a],\ldots)$. The transition rule requires $\xi = (\ldots,a,\ [b^1,b^2,\ldots,b^2,a^2],\ldots)$; so if $\xi = (\ldots,a,\ [b^1,b_{i_1}{}^2,\ldots,b_{i_n}{}^2,a^2],\ldots)$, set $\theta\xi = (\ldots,A,\ [B_{i1}, B_{i_2},\ldots,B_{i_n},B,A],\ldots)$. Yet to be assigned in these blocks are lower indices on b^1 and a^2 for which there are rq choices and indices on B, A for which there are $p(s - 1)$

$+ p - 1$ choices. In both cases the number of choices is identical so that a one-to-one assignment of indices within such blocks can be made. Still unassigned are blocks of a which separate the b-blocks, e.g., $\xi = (\ldots,a^2, [a,\ldots,a], b^1,\ldots)$. The transition rule requires $\xi = (\ldots,a^2, [a^1,\ldots,a^1], b^1,\ldots)$; so if $\xi = (\ldots,a^2, [a_{j_1}{}^1,\ldots, a_{j_n}{}^1],b^1,\ldots)$, set $\theta\xi = (\ldots,A, [A_{j_1},\ldots,A_{j_n}],B,\ldots)$. Now θ is defined a.e.; it is one-to-one, and it is measurable. From this definition it follows that $\theta\sigma = \sigma'\theta$; and finally since $h_{\theta\mu}(\sigma') = h_{\mu'}(\sigma')$, the previous theorem implies $\theta\mu = \mu'$.

Another case arises in which it is necessary to consider an alphabet $\mathfrak{A}' = \{1,\ldots, N + 1\}$ and a transition matrix T' defined by: $t'_{N+1,j} = 0, 1 < j \leq N + 1$: $t_{ij}' = 1$, otherwise. The characteristic equation of T' is $y^{N-1}(y^2 - Ny - 1) = 0$. By similar methods the following is proved.

THEOREM. *If $ps - qr = -1$ and $p + s = N$, then (Ξ,μ,σ) is conjugate to (Ξ',μ',σ') where μ and μ' are measures of maximal entropy.*

5. *Proof of Main Theorem.*—On the basis of §3 and §4, the proof of the main theorem will now be sketched. Given two automorphisms φ and φ' of the two-torus X which have the same entropy, it follows that det Φ = det Φ'. We first take up the case when det $\Phi = +1$. Let det $(\Phi - yI) = y^2 \pm Ny + 1$ and λ be the characteristic value of Φ of modulus greater than one. Elementary geometric considerations reveal that the torus can be partitioned into two parallelograms R_1 and R_2 whose sides consist of two connected segments through the origin in each of the two characteristic directions of Φ. Taking γ to be the partition into parallelograms determined by $R_i \cap \varphi R_j$, $i, j = 1, 2$, one sees that the associated transition matrix T has the structure of the matrices considered in §4. In fact, the letters a, b describe whether $C_i \in \gamma$ is contained in R_1 or R_2, and the superscript indicates whether C_i arose by intersecting with φR_1 or φR_2. Finally, letting ℓ_i denote the length of the side of C_i in the direction associated with λ and examining what happens to C_i under the action of φ, we see that $|\lambda| \ell_i = \sum_j t_{ij}\ell_j$; hence $|\lambda|$ is the largest positive characteristic value of T.

To complete the proof of the theorem, observe that by its very construction, and the fact that $\lambda \neq 1$, γ is a generator. Thus the corollary of §3 applies to φ, γ and analogously to φ', γ'. Taking this in conjunction with the first theorem of §4, the proof is done. Similar considerations apply in the case det $\Phi = -1$, and hence the main theorem has been proved.

* The results reported in this paper were obtained in the course of research jointly sponsored by the U. S. Air Force Office of Scientific Research (contract AF 49 (638)-1682) and by IBM.

[1] Blum, J. R., and D. L. Hanson, "On the isomorphism problem for Bernoulli schemes," *Bull. Am. Math. Soc.*, **69**, 221–223 (1963); Meshalkin, L. D., "A case of isomorphism of Bernoulli schemes," *Dokl. Akad. Nauk SSSR*, **128**, 41–44 (1959); Sinai, Ja. G., "Weak isomorphism of transformations with invariant measure," *Mat. Sb.*, **63** (105), 23–42 (1964), *Am. Math. Soc. Transl., Ser.* 2, **57**, 123–143 (1966).

[2] Adler, R. L., and R. Palais, "Homeomorphic conjugacy of automorphisms on the torus," *Proc. Am. Math. Soc.*, **16**, 1222–1225 (1965); Arov, D. Z., "Topological similitude of automorphisms and translations of compact commutative groups," *Uspehi Mat. Nauk*, **18**, 133–138 (1963).

[3] Arov, D. Z., "The computation of entropy for one class of group endomorphisms," *Zap. Mat. Otd. Fiz.-Mat. Fak. i Har'kov. Mat. Obsc.*, **30** (4), 48–69 (1964); Genis, A. L., "Metric properties of the endomorphisms of the n-dimensional torus," *Dokl. Akad. Nauk SSSR*, **138**, 991–993 (1961); *Sov. Math. Dokl.*, **2**, 750–752 (1961); Sinai, Ja. G., "On the concept of entropy of a dynamical system," *Dokl. Akad. Nauk SSSR*, **124**, 768–771 (1959).

[4] For a discussion of some of the material of this section from another point of view, see Parry, W., "Intrinsic Markov chains," *Trans. Am. Math. Soc.*, **112**, 55–66 (1964).

[5] Gantmacher, F. R., *The Theory of Matrices* (New York: Chelsea Publishing Co., 1959), vol. 2, chap. 13.