ENTROPY AND DENSITY

GEORGE H. STEIN

ABSTRACT. We prove that for any r, "entropy $=r$" is a dense
condition in the uniform topology.

In 1959, Rohlin [5] proved the set of transformations with zero
entropy is a dense G_δ in the uniform topology. See Parry [3, p. 103]
for a proof. In this note, we prove:

THEOREM. *For each* r, $\infty \geq r > 0$, *the set of transformations with
entropy equal to r is dense in the uniform topology.*

For the weak topology, the analogous theorem can be proved easily
with the aid of the conjugacy lemma (Halmos [2, pp. 77–78]): the
conjugacy class of an antiperiodic transformation is dense. To com-
plete the proof requires only the existence of antiperiodic trans-
formations with arbitrary entropy (Lemma 1, below). The conjugacy
class of an antiperiodic transformation is never dense in the uniform
topology. However, a simple lemma enabling one to compute the
entropy of a transformation in terms of its entropy on invariant pieces
provides the key.

More precisely, we consider the unit square with Lebesgue measure;
transformations which are bijections, measurable together with their
inverses, and measure-preserving. (The results are true for Lebesgue
spaces with nonatomic measure, but we choose the unit square for
ease in picturing a certain construction.) The uniform topology on
transformations is given by the metric: $d(S, T) = m(\{x : Sx \neq Tx\})$.
See [2, pp. 61–68] for more information.

COMPUTATION LEMMA. *Let* $T: I^2 \to I^2$ *be a transformation with*
$\{M, N\}$ *a measurable partition, each element of which is invariant. Then*

$$h(T) = h(T_m) \mid + h(T \mid_N).$$

Notation. $\mid$ denotes restriction, capital Greek is used for partitions,
$\bigvee$ denotes smallest common refinement, h is the entropy function,
etc. as in Billingsley [1].

REMARK. This follows from the more general formula of Rohlin
[4, p. 28]. However, a direct proof is easy.

PROOF. $h(T) = \sup_\Gamma h(\Gamma, T)$ by definition. If $\Gamma \subset \Lambda$, $h(\Gamma, T) \leqq h(\Lambda, T)$. So, noting that $\Gamma \subset \Gamma \vee \{M, N\}$, without loss of generality, we may assume that all partitions refine $\{M, N\}$. Now

$$h(\Gamma, T) = \lim_{n \to \infty} \frac{1}{n} H(\Gamma \vee \cdots \vee T^{n-1}\Gamma)$$

$$= \lim_{n \to \infty} \frac{-1}{n} \sum_{A \in \Gamma \vee \ldots \vee T^{n-1}\Gamma} m(A) \log(m(A))$$

$$= \lim_n \frac{-1}{n} \sum_{A \in (\Gamma \vee \ldots \vee T^{n-1}\Gamma) \cap M} m(A) \log(m(A))$$

$$+ \lim_n \frac{-1}{n} \sum_{A \in (\Gamma \vee \ldots \vee T^{n-1}\Gamma) \cap N} m(A) \log(m(A))$$

$$= \lim_n \frac{1}{n} H(\Gamma \cap M \vee \cdots \vee T\big|_M^{n-1}(\Gamma \cap M))$$

$$+ \lim_n \frac{1}{n} H(\Gamma \cap N \vee \cdots \vee T\big|_N^{n-1}(\Gamma \cap N))$$

$$\text{(since } (\Gamma \vee \cdots \vee T^{n-1}\Gamma) \cap M = (\Gamma \cap M) \vee$$

$$\cdots \vee T\big|_M^{n-1}(\Gamma \cap M) \text{ by invariance)}$$

$$= h(\Gamma \cap M, T\big|_M) + h(\Gamma \cap N, T\big|_N).$$

As Γ ranges over refinements of $\{M, N\}$, $\Gamma \cap N$ and $\Gamma \cap M$ range over partitions of N and M respectively. Hence, by positivity,

$$h(T) = \sup_\Gamma h(\Gamma, T) = \sup_\Gamma \{h(\Gamma \cap M, T\big|_M) + h(\Gamma \cap N, T\big|_N)\}$$

$$= \sup_\Lambda h(\Lambda, T\big|_M) + \sup_\Theta h(\Theta, T\big|_N) = h(T\big|_M) + h(T\big|_N). \quad \text{Q.E.D.}$$

COROLLARY. *The analogous formula is true when the elementwise invariant partition is countable.*

This follows from the more powerful theory developed in [3, Chapter 2] which deals with the difficulties that countable partitions raise. Every step follows exactly as above.

LEMMA 1. *There is an antiperiodic transformation on the unit square with any prescribed entropy.*

PROOF. Arbitrary Bernoulli shifts can be made conjugate to transformations on the unit square precisely in the same manner as

the shift $B(1/2, 1/2)$ is made conjugate to the Baker's transformation on the unit square. See [1] and [2] for more on Bernoulli shifts. Since conjugation preserves entropy [1, pp. 64–65], these transformations have entropy $-\sum_{i=1}^{i=n} p_i \log p_i$. Given $r > 0$, there is an n, and $(p_i)_{i=1}^{i=n}$ such that $-\sum_{i=1}^{i=n} p_i \log p_i = r$. This follows from the intermediate value theorem and the fact that, for fixed n, the function

$$f(p_1, \cdots, p_n) = -\sum_{i=1}^{i=n} p_i \log p_i$$

subject to the constraints: $p_i \geqq 0, \sum p_i = 1$, has a maximum equal to $\log n \to \infty$ as $n \to \infty$. For $r = \infty$, we conjugate the shift on countably many symbols,

$$B(1/2, 1/4, 1/16, \cdots, 1/2^{2n}, \cdots, 1/2^{2n}, \cdots),$$

$$2^{2n-n-1} \text{ times}$$

to a transformation in the unit square.

All Bernoulli shifts are antiperiodic.   Q.E.D.

The following lemma is used tacitly hereafter.

LEMMA 2. *If $E \subset I^2$ has positive measure, then E is isomorphic to I^2 in the sense that there is a measurable, invertible measure-preserving bijection which preserves the normalized inherited measure.*

This follows easily from Halmos' lemma [2, p. 74].

Notation. Let T be a transformation. Define

$$E_n = \{x : T^n x = x; T^j x \neq x, 0 \leq j < n\};$$

$E_\infty = \{x : T^n x \neq x \text{ for all } n\}$. These form a countable measurable partition each element of which is T-invariant. If $\{A, B\}$ is a measurable partition, define $d_A(S, T) = m(\{x \in A : Sx \neq Tx\})$. Then $d(S, T) = d_A(S, T) + d_B(S, T)$.

PROOF OF THE THEOREM. Let T be a transformation. We distinguish two cases: either

(i) almost every point is periodic; or

(ii) the set of nonperiodic points has positive measure.

(i) By the corollary to the computation lemma, and the fact that periodic transformations have zero entropy, $h(T) = 0$. Hence it suffices to find an invariant set of small measure: on this we may construct a transformation of arbitrary entropy by Lemma 1, and then apply the computation lemma. For some n, $m(E_n) > 0$. Find a set $Y \subset E_n$ such that $\epsilon/n > m(Y) > 0$. Then $Y \cup \cdots \cup T^{n-1} Y$ is invariant and of small enough measure.

(ii) We reduce case (ii) to case (i) by use of the following fact: arbitrarily near any transformation T is an S which has almost every point periodic. The uniform approximation theorem of Halmos [2, p. 75] applied to $T|_{E_\infty}: E_\infty \to E_\infty$, yields, for any n and ϵ, a $T': E_\infty \to E_\infty$ which has period n and is within $1/n + \epsilon$ of T. Choosing n large, defining S to be T on $I^2 - E_\infty$ and to be T' on E_∞, we have

$$d(S, T) = d_{E_\infty}(S, T) + d_{I^2 - E_\infty}(S, T) = 1/n + \epsilon + 0 < 2\epsilon. \quad \text{Q.E.D.}$$

REMARK. From (ii) above it follows easily that periodic transformations are dense, a result used in [3, p. 103].

PROOF. Take S as above in (ii). If S is periodic, we are done. If not, choose N so that $\epsilon > m(\bigcup_{n > N} E_n) > 0$. Then $S|\bigcup_{0 \le n \le N} E_n$ is periodic of period at most $N!$ On $\bigcup_{n > N} E_n$, modify S to be periodic with period $N!$ The resulting transformation is periodic and within 3ϵ of an arbitrary T.

REFERENCES

1. Patrick Billingsley, *Ergodic theory and information*, Wiley, New York, 1965. MR **33** #254.

2. P. R. Halmos, *Lectures on ergodic theory*, Chelsea, New York, 1956. MR **20** #3958.

3. William Parry, *Entropy and generators in ergodic theory*, Benjamin, New York, 1969.

4. V. A. Rohlin, *Lectures on entropy theory of transformations with invariant measure*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 3–56 = Russian Math. Surveys **22** (1967), no. 5, 1–52. MR **36** #349.

5. ———, *Entropy of metric automorphism*, Dokl. Akad. Nauk SSSR **124** (1959), 980–983. (Russian) MR **21** #2037; errata, MR **21** p. 1599.

COLUMBIA UNIVERSITY, NEW YORK, NEW YORK 10027