THE ENTROPY OF CHEBYSHEV POLYNOMIALS

BY

R. L. ADLER AND M. H. McANDREW

1. **Introduction.** The purpose of this work is to compute the topological entropy of the νth Chebyshev polynomial $T_\nu(x)$ considered as a map of $[-1,1]$ onto itself. The notation and basic definitions relevant to the concept of topological entropy are contained in [1] and are reviewed briefly below.

For an open cover $\mathfrak{A}$ of a compact space $X, N(\mathfrak{A})$ denotes the minimum cardinality of all sub-covers of $\mathfrak{A}$. $H(\mathfrak{A}) = \log N(\mathfrak{A})$ is called the *entropy* of $\mathfrak{A}$. A cover $\mathfrak{B}$ is said to *refine* a cover $\mathfrak{A}$ if every set of $\mathfrak{B}$ is a subset of some set of $\mathfrak{A}$; we use the notation $\mathfrak{A} \prec \mathfrak{B}$. We define the *join* of two covers $\mathfrak{A}, \mathfrak{B}$ to be the cover $\mathfrak{A} \vee \mathfrak{B} = \{A \cap B;\ A \in \mathfrak{A},\ B \in \mathfrak{B}\}$. For a continuous map ϕ of X into itself we define $h(\phi, \mathfrak{A})$, the *entropy of ϕ with respect to* $\mathfrak{A}$ to be

$$\lim_{n \to \infty} H(\mathfrak{A} \vee \phi^{-1}\mathfrak{A} \vee \cdots \vee \phi^{-n+1}\mathfrak{A})/n;$$

in [1] this limit is shown to exist. Finally $h(\phi)$, the *entropy of ϕ*, is defined to be $\sup h(\phi, \mathfrak{A})$ where the supremum is taken over all open covers $\mathfrak{A}$ of X. In the sequel we use the following properties.

(1) $\prec$ is transitive.

(2) $\mathfrak{A} \prec \mathfrak{A}'$ and $\mathfrak{B} \prec \mathfrak{B}' \Rightarrow \mathfrak{A} \vee \mathfrak{B} \prec \mathfrak{A}' \vee \mathfrak{B}'$.

(3) $\mathfrak{A} \prec \mathfrak{B} \Rightarrow N(\mathfrak{A}) \leqq N(\mathfrak{B})$.

(4) $\mathfrak{A} \prec \mathfrak{B} \Rightarrow \phi^{-1}\mathfrak{A} \prec \phi^{-1}\mathfrak{B}$.

(5) $\phi^{-1}(\mathfrak{A} \vee \mathfrak{B}) = \phi^{-1}\mathfrak{A} \vee \phi^{-1}\mathfrak{B}$.

(6) Let $\mathfrak{A}_n$ be a refining sequence; i.e. a sequence of open covers such that $\mathfrak{A}_n \prec \mathfrak{A}_{n+1}$ and for every open cover $\mathfrak{B}$ there is some $\mathfrak{A}_n$ with $\mathfrak{B} \prec \mathfrak{A}_n$. Then $h(\phi) = \lim_{n \to \infty} h(\phi, \mathfrak{A}_n)$. These properties are proved in [1].

2. **Preliminary lemmas.**

LEMMA 1. *Let X be a compact topological space and μ a Borel measure on X. For an open cover $\mathfrak{B}$ of X, let $g(\mathfrak{B}, x) = 1/\sup \mu(B)$, the supremum being taken over all B with $x \in B$ and $B \in \mathfrak{B}$. Then $\int_X g(\mathfrak{B}, x) d\mu \leqq N(\mathfrak{B})$.*

Proof. $g(\mathfrak{B}, x)$ is measurable since $\{x : g(\mathfrak{B}, x) < \lambda\} = \bigcup_{\mu(B_i) > 1/\lambda} B_i$, an open set.

Let $\mathfrak{B}' = \{B_1, B_2, \cdots, B_{N(\mathfrak{B})}\}$ be a subcover of minimal cardinality. For $x \in X$ let $B(x)$ be that B_i of least index such that $x \in B_i$. Then $\{x : B(x) = B_i\}$ is just

Received by the editors March 15, 1965

$B_i \cap \bar{B}_1 \cap \bar{B}_2 \cap \cdots \cap \bar{B}_{i-1}$ and is measurable. If $\mu(B_i) = 0$ then $\mu\{x : B(x) = B_i\} = 0$ and

$$\int_{\{x : B(x) = B_i\}} g(\mathfrak{B}, x) d\mu = 0.$$

If $\mu(B_i) \neq 0$ then

$$\int_{\{x : B(x) = B_1\}} g(\mathfrak{B}, x) d\mu \leqq \int_{B_i} \frac{1}{\mu(B_i)} d\mu = 1.$$

Since $X = \bigcup_{i=1}^{N(\mathfrak{B})} \{x : B(x) = B_i\}$ the result of the lemma now follows.

Lemma 2. *Let $v \geqq 2$. Then there is a function $\lambda(r)$ defined for integral $r \geqq 2$ with the following properties*:

(2.1) (i) $\lim_{r \to \infty} \lambda(r) = v$.

(ii) *If $r \geqq 2$ and $\{I_n ; n \geqq 0\}$ is a sequence of real numbers satisfying*

$$I_{n+1} > vI_n - (v-1) I_{n-s+1} \tag{2.2}$$

for $1 \leqq s \leqq r$ and $s \leqq n+1$, then

$$\liminf_{n \to \infty} I_n^{1/n} \geqq \lambda(r). \tag{2.3}$$

Proof. We shall show that the unique positive zero of

$$f_r(x) = x^{r-1} - (v-1)(x^{r-2} + x^{r-3} + \cdots + 1) \tag{2.4}$$

has the properties required for $\lambda(r)$. We note that for $r > 2$, $\lambda(r)$ is the positive zero other than 1 of $g_r(x) = (x-1) f_r(x) = x^r - vx^{r-1} + v - 1$. Now $g_r(v) = v - 1 > 0$, and $g_r(v - v^{2-r}) = v - 1 - v(1 - v^{1-r})^{r-1}$. Clearly $g_r(v - v^{2-r}) \to -1$ as $r \to \infty$. Hence for r sufficiently large, $v - v^{2-r} < \lambda(r) < v$. This verifies (2.1).

To verify the second property of $\lambda(r)$ let $r \geqq 2$ and let I_n be a sequence satisfying (2.2). Let $J_n = I_{n+1} - I_n$ $(n \geqq 0)$. Then from (2.2) with $s = 1$, $J_n > 0$. Further

$$J_n > (v-1)(J_{n-1} + J_{n-2} + \cdots + J_{n-s+1}) \tag{2.5}$$

for $2 \leqq s \leqq r$ and $s \leqq n+1$. We shall show that for $n \geqq 0$,

$$J_n \geqq J_0 \lambda(r)^{n-r}. \tag{2.6}$$

Since $f_r(1) = 1 - (v-1)(r-1) \leqq 0$ and $f_r(+\infty) = +\infty$, $\lambda(r) \geqq 1$. Hence (2.6) is true for $n = 0$. From (2.5) with $s = 2, 3, \cdots, r-1$ and $n = s-1$ it follows that $J_n > J_0$ for $1 \leqq n \leqq r-2$ and, a fortiori, (2.6) is true. The remaining cases for n follow from (2.5) with $s = r$ by induction, since $\lambda(r)$ is a zero of (2.4). Relation (2.3) now follows from (2.6) since $I_n = I_1 + \sum_{m=1}^{n-1} J_m$.

3. **Main result.** Let X be the interval $[-1, 1]$ with its usual topology and let ϕ be the map $x \to T_\nu(x)$ where T_ν is the νth Chebyshev polynomial; i.e. $T_\nu(\cos\theta) = \cos\nu\theta$.

THEOREM. $h(\phi) = \log \nu$.

Proof. Let $X' = [0, \pi]$ and let σ be the homeomorphism of X onto X' defined by $x' = \sigma x = \cos^{-1}x$. Let ψ be the continuous map $\sigma\phi\sigma^{-1}$ of X' onto X'. By Theorem 1 of [1], $h(\psi) = h(\phi)$ and so we may work with ψ instead of ϕ. The map ψ is given explicitly by $\psi(x') = S_\nu(x')$ where

$$S_\nu(x) = \begin{cases} \nu x - k\pi, & k \text{ even}, \\ (k+1)\pi - \nu x, & k \text{ odd}, \end{cases}$$

for $k\pi/\nu \leqq x \leqq (k+1)\pi/\nu$, $k = 0, 1, \cdots, \nu - 1$. Figure 1 illustrates the case $\nu = 3$. Now S_1 is just the identity transformation on X' and hence for $\nu = 1$, $h(\psi) = 0$.

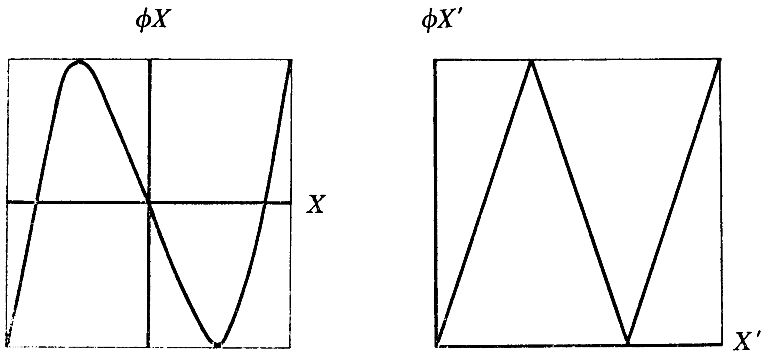

FIGURE 1

For $\nu > 1$, we argue as follows. Let $\varepsilon < 1$ and let $\mathfrak{A}_\varepsilon$ be the cover of X' consisting of all intervals of length $\leqq \varepsilon$ of the type $(a, b,)$, $[0, b)$ or $(a, \pi]$. For such an interval I of length l, $\psi^{-1}I$ is the union of disjoint similar intervals each of length l' where $l/\nu \leqq l' \leqq 2l/\nu$; this is clear from Figure 2. Hence $\psi^{-1}\mathfrak{A}_\varepsilon \prec \mathfrak{A}_{\varepsilon/\nu}$. By properties (1) and (4) of the introduction it follows that $\psi^{-k}\mathfrak{A}_\varepsilon \prec \mathfrak{A}_{\varepsilon/\nu^k}$ for $k = 1, 2,$ Hence

$$\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \cdots \vee \psi^{-n}\mathfrak{A}_\varepsilon \prec \mathfrak{A}_\varepsilon \vee \mathfrak{A}_{\varepsilon/\nu} \vee \cdots \vee \mathfrak{A}_{\varepsilon/\nu^n} = \mathfrak{A}_{\varepsilon/\nu^n},$$

since $\mathfrak{A}_{\varepsilon/\nu}r \prec \mathfrak{A}_{\varepsilon/\nu}n$ for $0 \leqq r \leqq n$. Therefore, by property (3),

$$N(\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \cdots \vee \psi^{-n}\mathfrak{A}_\varepsilon) \leqq N(\mathfrak{A}_{\varepsilon/\nu^n}) \leqq \pi\nu^n/\varepsilon + 1.$$

Therefore $h(\psi, \mathfrak{A}_\varepsilon) \leqq \log \nu$. Now the sequence $\{\mathfrak{A}_{1/n}\}$ is shown in [1] to be a refining sequence and so, by property (6), $h(\psi) = \lim_{n\to\infty} h(\psi, \mathfrak{A}_{1/n}) \leqq \log \nu$.

Next we will prove the reverse inequality, $h(\psi) \geqq \log v$. Let μ be Lebesgue measure on X' and let $g(\mathfrak{B}, x)$ be defined for $x \in X'$ and $\mathfrak{B}$ a cover of X', as in Lemma 1. Suppose now that $\varepsilon < \pi/2v$. We note first that if $\mathfrak{B}$ is an open cover whose sets have diameter $< v\varepsilon$,

$$g(\psi^{-} \mathfrak{B} \vee \mathfrak{A}_{\varepsilon}, x) \geqq g(\mathfrak{B}, \psi x), \tag{3.1}$$

and, for $x \notin S_{\varepsilon}$,

$$g(\psi^{-1} \mathfrak{B} \vee \mathfrak{A}_{\varepsilon}, x) = v g(\mathfrak{B}, \psi x), \tag{3.2}$$

where S_{ε} is the set of points at distance $\leqq \varepsilon$ from some $\pi k/v$ with $0 < k < v$. The proof of (3.1) and (3.2) is immediate from Figure 2, where B represents some set of $\mathfrak{B}$. Inequality (3.1) follows since $\mu(\psi^{-1}B) = \mu(B)$ and $\mu(\psi^{-1}B \cap A) \leqq \mu(\psi^{-1}B)$ for any A. For (3.2), the essential point is that for any $B \in \mathfrak{B}, \psi^{-1}B$ consists of exactly v pieces of each measure $\mu(B)/v$ and di-

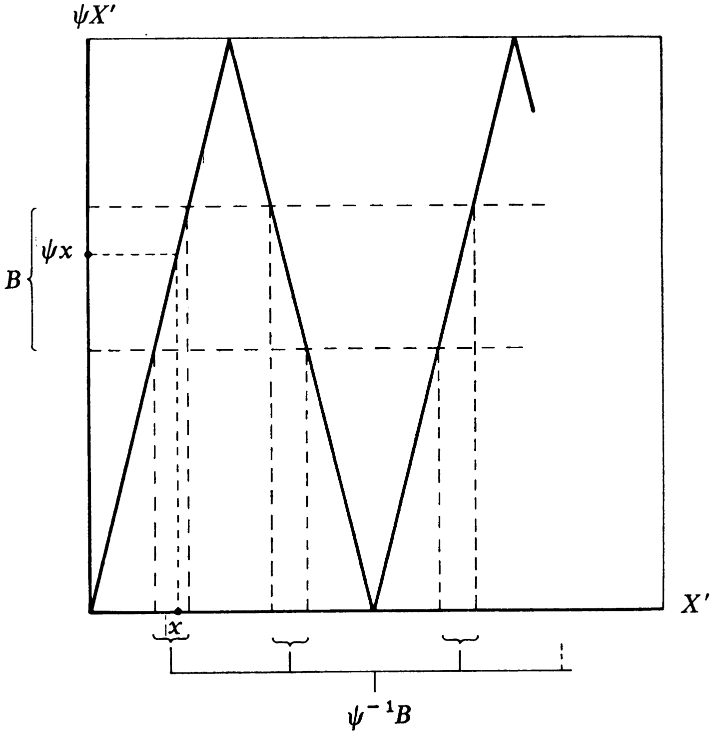

FIGURE 2

ameter $d(B)/v$ where $d(B)$ is the diameter of B. If $x \notin S_{\varepsilon}$ and $x \in A \in \mathfrak{A}_{\varepsilon}$ then $A \cap \psi^{-1}B$ contains points of at most one such piece and there is a choice of

$A \in \mathfrak{A}_\varepsilon$ such that $A \cap \psi^{-1}B$ is the whole of one piece. Let $g_n(x)$ denote $g(\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \cdots \vee \psi^{-n}\mathfrak{A}_\varepsilon, x)$. Taking $\mathfrak{B} = \mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \cdots \vee \psi^{-n}\mathfrak{A}_\varepsilon$ in (3.1) and (3.2), we obtain

$$g_{n+1}(x) \geqq g_n(\psi x), \tag{3.3}$$

and, for $x \notin S_\varepsilon$,

$$g_{n+1}(x) = v g_n(\psi x). \tag{3.4}$$

From (3.3) and (3.4) we have, for $0 \leqq k \leqq v-1$,

$$\int_{k\pi/v}^{(k+1)\pi/v} g_{n+1}(x)dx \geqq \int_{k\pi/v+\varepsilon}^{(k+1)\pi/v-\varepsilon} v g_n(\psi x)dx + \int_{k\pi/v}^{k\pi/v+\varepsilon} g_n(\psi x)dx + \int_{(k+1)\pi/v-\varepsilon}^{(k+1)\pi/v} g_n(\psi x)dx$$

$$= \int_{v\varepsilon}^{\pi-v\varepsilon} g_n(y)dy + v^{-1}\int_0^{v\varepsilon} g_n(y)dy + v^{-1}\int_{\pi-v\varepsilon}^{\pi} g_n(y)dy.$$

Hence

$$\int_0^\pi g_{n+1}(x)dx \geqq v\int_0^\pi g_n(y)dy - (v-1)\int_0^{v\varepsilon} g_n(y)dy - (v-1)\int_{\pi-v\varepsilon}^{\pi} g_n(y)dy. \tag{3.5}$$

Now for $0 < a < \pi/v - \varepsilon$, $[0,a] \cap S_\varepsilon = \phi$. Hence for $n \geqq 1$,

$$\int_0^a g_n(x)dx = \int_0^a v g_{n-1}(\psi x)dx = \int_0^{va} g_{n-1}(y)dy.$$

Iterating this operation we have that if

$$0 < v^{r-1}a < \pi/v - \varepsilon \text{ and } n \geqq r \geqq 0, \tag{3.6}$$

then

$$\int_0^a g_n(x)dx = \int_0^{av^r} g_{n-r}(y)dy. \tag{3.7}$$

Similarly if v is odd (so that $\psi(\pi) = \pi$) and a, n, r satisfy (3.6) then

$$\int_{\pi-a}^{\pi} g_n(x)dx = \int_{\pi-av^r}^{\pi} g_{n-r}(x)dx. \tag{3.8}$$

Further (3.8) also holds if v is even and a, n, r satisfy (3.6). In this case $\psi(x)$ is an even function of $x - \pi/2$ and cover $\mathfrak{A}_\varepsilon$ is symmetric about $\pi/2$; hence $g_n(x)$ is an even function of $x - \pi/2$ and now the left-hand and right-hand sides of (3.7) and (3.8) are respectively equal.

Let $I_n = \int_0^\pi g_n(x)dx$ and choose $r_0 = r_0(\varepsilon)$ such that $0 < \varepsilon v^{r_0-1} < \pi/v - \varepsilon$ and $2v^{r_0}\varepsilon < \pi/v$. Let $1 \leqq s \leqq r_0$ and $n \geqq s-1$. Then from (3.5),

$$(3.9)\quad \begin{aligned} I_{n+1} &\geqq vI_n - (v-1)\int_0^{v\varepsilon} g_n(x)dx - (v-1)\int_{\pi - v\varepsilon}^{\pi} g_n(x)dx, \\ &= vI_n - (v-1)\int_0^{v^s\varepsilon} g_{n-s+1}(y)dy - (v-1)\int_{\pi - v^s\varepsilon}^{\pi} g_{n-s+1}(y)dy, \end{aligned}$$

from (3.7) and (3.8) with $a = v\varepsilon$ and $r = s-1$. By definition of r_0, $v^s\varepsilon < \pi - v^s\varepsilon$, and clearly $g_n(x) > 0$ for all $0 \leqq x \leqq \pi$. Hence from (3.9),

$$I_{n+1} > vI_n - (v-1)I_{n-s+1}$$

for $1 \leqq s \leqq r_0$ and $n \geqq s-1$. Let $\lambda(r)$ be defined as in Lemma 2. Then

$$\begin{aligned} h(\psi, \mathfrak{A}_\varepsilon) &= \log\left(\lim_{n\to\infty} N^{1/n}(\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \cdots \vee \psi^{-n+1}\mathfrak{A}_\varepsilon)\right) \\ &\geqq \log\left(\liminf_{n\to\infty} I_{n-1}^{1/n}\right), \text{ by Lemma 1,} \\ &\geqq \log \lambda(r_0), \qquad \text{by Lemma 2.} \end{aligned}$$

Letting $\varepsilon \to 0$ we may choose $r_0(\varepsilon)$ so that $r_0 \to \infty$ and hence $h(\psi) \geqq \sup_\varepsilon h(\psi, \mathfrak{A}_\varepsilon) \geqq \log v$ since $\lim_{r\to\infty} \lambda(r_0) = v$. This concludes the proof of the theorem.

References

1. R. L. Adler, A. G. Konheim and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319.

International Business Machines Corporation,
Yorktown Heights, New York