

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 n th Chebyshev polynomial $T_n(x)$ 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

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

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) $\mathfrak{A} \prec \mathfrak{B} \Rightarrow N(\mathfrak{A}) \leq N(\mathfrak{B})$.
- (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 $h(\phi) = \lim_{n \rightarrow \infty} h(\phi, \mathfrak{A}_n)$. 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 $\int_X g(\mathfrak{B}, x) d\mu \leq N(\mathfrak{B})$.*

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 $\mathfrak{B}' = \{B_1, B_2, \dots, B_{N(\mathfrak{B})}\}$ 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

$B_i \cap \bar{B}_1 \cap \bar{B}_2 \cap \dots \cap \bar{B}_{i-1}$ 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

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

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

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

- (2.1) (i) $\lim_{r \rightarrow \infty} \lambda(r) = v$.
- (ii) *If $r \geq 2$ and $\{I_n; n \geq 0\}$ is a sequence of real numbers satisfying*

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

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

$$(2.3) \quad \liminf_{n \rightarrow \infty} I_n^{1/n} \geq \lambda(r).$$

Proof. We shall show that the unique positive zero of

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

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 $g_r(v-v^{2-r}) \rightarrow -1$ as $r \rightarrow \infty$. 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 $r \geq 2$ and let I_n be a sequence satisfying (2.2). Let $J_n = I_{n+1} - I_n$ ($n \geq 0$). Then from (2.2) with $s = 1$, $J_n > 0$. Further

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

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

$$(2.6) \quad J_n \geq J_0 \lambda(r)^{n-r}.$$

Since $f_r(1) = 1 - (v-1)(r-1) \leq 0$ and $f_r(+\infty) = +\infty$, $\lambda(r) \geq 1$. Hence (2.6) is true for $n = 0$. From (2.5) with $s = 2, 3, \dots, r-1$ and $n = s-1$ it follows that $J_n > J_0$ for $1 \leq n \leq r-2$ 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 $x \rightarrow T_v(x)$ where T_v is the v th Chebyshev polynomial; i.e. $T_v(\cos \theta) = \cos v\theta$.

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

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 $\psi(x') = S_v(x')$ where

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

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

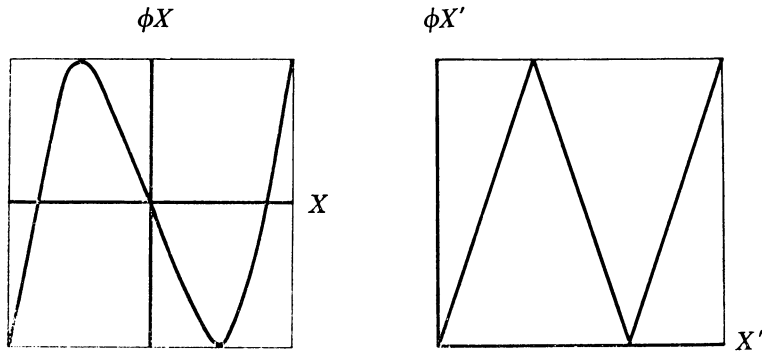


FIGURE 1

For $v > 1$, we argue as follows. Let $\varepsilon < 1$ and let \mathfrak{A}_ε be the cover of X' consisting of all intervals of length $\leq \varepsilon$ 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 $l/v \leq l' \leq 2l/v$; this is clear from Figure 2. Hence $\psi^{-1}\mathfrak{A}_\varepsilon \prec \mathfrak{A}_{\varepsilon/v}$. By properties (1) and (4) of the introduction it follows that $\psi^{-k}\mathfrak{A}_\varepsilon \prec \mathfrak{A}_{\varepsilon/v^k}$ for $k = 1, 2$, Hence

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

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

$$N(\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \dots \vee \psi^{-n}\mathfrak{A}_\varepsilon) \leq N(\mathfrak{A}_{\varepsilon/v^n}) \leq \pi v^n / \varepsilon + 1.$$

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

Next we will prove the reverse inequality, $h(\psi) \geq \log v$. 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 $< v\varepsilon$,

$$(3.1) \quad g(\psi^{-1} \mathfrak{B} \vee \mathfrak{A}_\varepsilon, x) \geq g(\mathfrak{B}, \psi x),$$

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

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

where S_ε is the set of points at distance $\leq \varepsilon$ 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 $\mu(\psi^{-1}B \cap A) \leq \mu(\psi^{-1}B)$ 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-

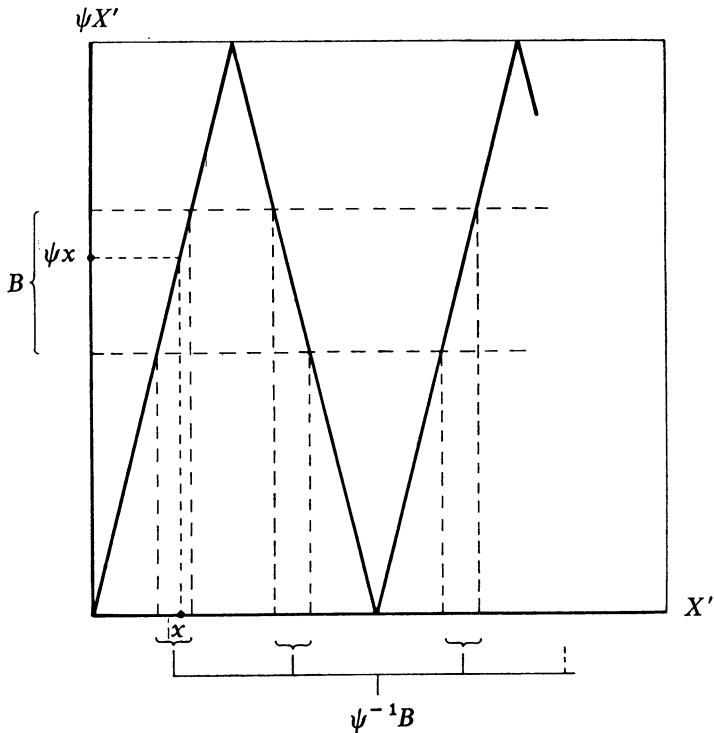


FIGURE 2

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 $g(\mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \dots \vee \psi^{-n}\mathfrak{A}_\varepsilon, x)$. Taking $\mathfrak{B} = \mathfrak{A}_\varepsilon \vee \psi^{-1}\mathfrak{A}_\varepsilon \vee \dots \vee \psi^{-n}\mathfrak{A}_\varepsilon$ in (3.1) and (3.2), we obtain

$$(3.3) \quad g_{n+1}(x) \geq g_n(\psi x),$$

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

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

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

$$\begin{aligned} \int_{k\pi/v}^{(k+1)\pi/v} g_{n+1}(x) dx &\geq \int_{k\pi/v+\varepsilon}^{(k+1)\pi/v-\varepsilon} v g_n(\psi x) dx + \int_{k\pi/v}^{k\pi/v+\varepsilon} g_n(\psi x) dx + \int_{(k+1)\pi/v-\varepsilon}^{(k+1)\pi/v} g_n(\psi x) dx \\ &= \int_{v\varepsilon}^{\pi-v\varepsilon} g_n(y) dy + v^{-1} \int_0^{v\varepsilon} g_n(y) dy + v^{-1} \int_{\pi-v\varepsilon}^\pi g_n(y) dy. \end{aligned}$$

Hence

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

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

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

Iterating this operation we have that if

$$(3.6) \quad 0 < v^{r-1}a < \pi/v - \varepsilon \text{ and } n \geq r \geq 0,$$

then

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

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

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

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}_ε 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 $1 \leq s \leq r_0$ and $n \geq s - 1$. Then from (3.5),

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

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 $0 \leq x \leq \pi$. Hence from (3.9),

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

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

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

Letting $\varepsilon \rightarrow 0$ we may choose $r_0(\varepsilon)$ so that $r_0 \rightarrow \infty$ and hence $h(\psi) \geq \sup_\varepsilon h(\psi, \mathfrak{A}_\varepsilon) \geq \log v$ since $\lim_{r \rightarrow \infty} \lambda(r) = v$. 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