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

\[
\lim_{n \to \infty} \frac{H(\mathcal{A} \vee \phi^{-1} \mathcal{A} \vee \cdots \vee \phi^{-n+1} \mathcal{A})}{n};
\]

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

(1) \( \prec \) is transitive.
(2) \( \mathcal{A} \prec \mathcal{A}' \) and \( \mathcal{B} \prec \mathcal{B}' \Rightarrow \mathcal{A} \vee \mathcal{B} \prec \mathcal{A}' \vee \mathcal{B}' \).
(3) \( \mathcal{A} \prec \mathcal{B} \Rightarrow N(\mathcal{A}) \leq N(\mathcal{B}) \).
(4) \( \mathcal{A} \prec \mathcal{B} \Rightarrow \phi^{-1} \mathcal{A} \prec \phi^{-1} \mathcal{B} \).
(5) \( \phi^{-1}(\mathcal{A} \vee \mathcal{B}) = \phi^{-1} \mathcal{A} \vee \phi^{-1} \mathcal{B} \).
(6) Let \( \mathcal{A}_n \) be a refining sequence; i.e. a sequence of open covers such that \( \mathcal{A}_n \prec \mathcal{A}_{n+1} \) and for every open cover \( \mathcal{B} \) there is some \( \mathcal{A}_n \) with \( \mathcal{B} \prec \mathcal{A}_n \). Then \( h(\phi) = \lim_{n \to \infty} h(\phi, \mathcal{A}_n) \). These properties are proved in [1].

2. Preliminary lemmas.

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

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

Let \( \mathcal{B}' = \{ B_1, B_2, \cdots, B_{N(\mathcal{B})} \} \) 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
THE ENTROPY OF CHEBYSHEV POLYNOMIALS

$B_1 \cap \bar{B}_1 \cap B_2 \cap \cdots \cap \bar{B}_{i-1}$ and is measurable. If $\mu(B_i) = 0$ then $\mu\{x: B(x) = B_i\} = 0$ and

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

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

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

Since $X = \bigcup_{i=1}^{N(\mathcal{B})} \{x: B(x) = B_i\}$ 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) $\lim_{r \to \infty} \lambda(r) = v$.

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

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

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

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

**Proof.** We shall show that the unique positive zero of

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

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 $g_r(v - v^{2-r}) = v - 1 - v(1 - v^{1-r}) - 1$. 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 $r \geq 2$ 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) $J_n > (v - 1)(J_{n-1} + J_{n-2} + \cdots + J_{n-r+1})$

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

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

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 $s = 2, 3, \cdots, r - 1$ 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 $I_n = I_1 + \sum_{m=1}^{n-1} J_m$.  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
3. Main result. Let \( X \) be the interval \([-1, 1]\) with its usual topology and let \( \phi \) be the map \( x \to T_v(x) \) where \( T_v \) is the \( v \)th Chebyshev polynomial; i.e. \( T_v(\cos \theta) = \cos v\theta \).

\textbf{Theorem.} \( h(\phi) = \log v \).

\textbf{Proof.} Let \( X' = [0, \pi] \) and let \( \sigma \) be the homeomorphism of \( X \) onto \( X' \) defined by \( x' = \sigma x = \cos^{-1}x \). Let \( \psi \) 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 \( \psi \) instead of \( \phi \). The map \( \psi \) is given explicitly by \( \psi(x') = S_v(x') \) where

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

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

For \( v > 1 \), we argue as follows. Let \( \epsilon < 1 \) and let \( \mathcal{U}_\epsilon \) be the cover of \( X' \) consisting of all intervals of length \( \leq \epsilon \) 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 \( \psi^{-1}\mathcal{U}_\epsilon \subset \mathcal{U}_{\epsilon/v} \). By properties (1) and (4) of the introduction it follows that \( \psi^{-k}\mathcal{U}_\epsilon \subset \mathcal{U}_{\epsilon/v^k} \) for \( k = 1, 2, \ldots \). Hence

\[
\mathcal{U}_\epsilon \vee \psi^{-1}\mathcal{U}_\epsilon \vee \cdots \vee \psi^{-n}\mathcal{U}_\epsilon \subset \mathcal{U}_\epsilon \vee \mathcal{U}_{\epsilon/v} \vee \cdots \vee \mathcal{U}_{\epsilon/v^n} = \mathcal{U}_{\epsilon/v^n},
\]

since \( \mathcal{U}_{\epsilon/v^r} \subset \mathcal{U}_{\epsilon/v^n} \) for \( 0 \leq r \leq n \). Therefore, by property (3),

\[
N(\mathcal{U}_\epsilon \vee \psi^{-1}\mathcal{U}_\epsilon \vee \cdots \vee \psi^{-n}\mathcal{U}_\epsilon) \leq N(\mathcal{U}_{\epsilon/v^n}) \leq \pi v^n/\epsilon + 1.
\]

Therefore \( h(\psi, \mathcal{U}_\epsilon) \leq \log v \). Now the sequence \( \{\mathcal{U}_{1/n}\} \) is shown in [1] to be a refining sequence and so, by property (6), \( h(\psi, \mathcal{U}_{1/n}) \leq \log v \).
Next we will prove the reverse inequality, \( h(\psi) \geq \log v \). Let \( \mu \) be Lebesgue measure on \( X' \) and let \( g(\mathcal{B}, x) \) be defined for \( x \in X' \) and \( \mathcal{B} \) a cover of \( X' \), as in Lemma 1. Suppose now that \( \varepsilon < \pi/2v \). We note first that if \( \mathcal{B} \) is an open cover whose sets have diameter < \( v\varepsilon \),

\[
(3.1) \quad g(\psi^{-1}\mathcal{B} \cup \mathcal{A}_\varepsilon, x) \geq g(\mathcal{B}, \psi x),
\]

and, for \( x \notin S_\varepsilon \),

\[
(3.2) \quad g(\psi^{-1}\mathcal{B} \cup \mathcal{A}_\varepsilon, x) = v g(\mathcal{B}, \psi x),
\]

where \( S_\varepsilon \) 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 \( \mathcal{B} \). 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 \( B \in \mathcal{B}, \psi^{-1}B \) consists of exactly \( v \) pieces of each measure \( \mu(B)/v \) and diameter \( d(B)/v \) where \( d(B) \) is the diameter of \( B \). If \( x \notin S_\varepsilon \) and \( x \in A \in \mathcal{A}_\varepsilon \) then \( A \cap \psi^{-1}B \) contains points of at most one such piece and there is a choice of

![Figure 2](image-url)
such that \( A \cap \psi^{-1}B \) is the whole of one piece. Let \( g_n(x) \) denote \( g(\mathcal{A}_e \cup \psi^{-1}\mathcal{A}_e \cup \cdots \cup \psi^{-n}\mathcal{A}_e, x) \). Taking \( \mathcal{B} = \mathcal{A}_e \cup \psi^{-1}\mathcal{A}_e \cup \cdots \cup \psi^{-n}\mathcal{A}_e \) in (3.1) and (3.2), we obtain

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

and, for \( x \notin S_e \),

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

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

\[
\int_{kn/v}^{(k+1)n/v} g_{n+1}(x) \, dx \geq \int_{kn/v}^{(k+1)n/v} v g_n(\psi x) \, dx + \int_{kn/v}^{(k+1)n/v} g_n(\psi x) \, dx + \int_{kn/v}^{(k+1)n/v} g_n(\psi x) \, dx
\]

\[
= \int_{v^{-1}n}^{n} g_n(y) \, dy + v^{-1} \int_{v^{-1}n}^{v^{-1}n} g_n(y) \, dy + v^{-1} \int_{v^{-1}n}^{0} g_n(y) \, dy.
\]

Hence

\[
(3.5) \quad \int_{0}^{n} g_{n+1}(x) \, dx \geq v \int_{0}^{n} g_n(y) \, dy - (v-1) \int_{0}^{v} g_n(y) \, dy - (v-1) \int_{v^{-1}n}^{0} g_n(y) \, dy.
\]

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

\[
\int_{0}^{a} g_n(x) \, dx = \int_{0}^{a} v g_{n-1}(\psi x) \, dx = \int_{0}^{v} g_{n-1}(y) \, dy.
\]

Iterating this operation we have that if

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

then

\[
(3.7) \quad \int_{0}^{a} g_n(x) \, dx = \int_{0}^{av} g_{n-r}(y) \, dy.
\]

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

\[
(3.8) \quad \int_{n-a}^{n} g_n(x) \, dx = \int_{n-ar}^{n} g_{n-r}(x) \, dx.
\]

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 \( \mathcal{A}_e \) 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 \( I_n = \int_{0}^{n} g_n(x) \, dx \) and choose \( r_0 = r_0(\varepsilon) \) such that \( 0 < v^{r_0} < \pi/v - \varepsilon \) and \( 2v^{r_0} < \pi/v \). Let \( 1 \leq s \leq r_0 \) and \( n \geq s - 1 \). Then from (3.5),
\[ I_{n+1} \geq vI_n - (v-1) \int_0^{v\epsilon} g_n(x) dx - (v-1) \int_{\pi - v\epsilon}^{\pi} g_n(x) dx, \]
\[ = vI_n - (v-1) \int_0^{v\epsilon} g_{n-s+1}(y) dy - (v-1) \int_{\pi - v\epsilon}^{\pi} g_{n-s+1}(y) dy, \]
from (3.7) and (3.8) with \( a = v\epsilon \) and \( r = s - 1 \). By definition of \( r_0, v^s \epsilon < \pi - v^s \epsilon, \) 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
\[ h(\psi, \mathcal{A}_e) = \log \left( \lim_{n \to \infty} N^{1/n}(\mathcal{A}_e \cup \psi^{-1}\mathcal{A}_e \cup \cdots \cup \psi^{-(n+1)}\mathcal{A}_e) \right) \]
\[ \geq \log \left( \lim \inf_{n \to \infty} I_{n-1}^{1/n} \right), \text{ by Lemma 1,} \]
\[ \geq \log \lambda(r_0), \text{ by Lemma 2.} \]

Letting \( \epsilon \to 0 \) we may choose \( r_0(\epsilon) \) so that \( r_0(\epsilon) \to \infty \) and hence \( h(\psi) \geq \sup \lambda \) \( h(\psi, \mathcal{A}_e) \geq \log v \) since \( \lim_{r \to \infty} \lambda(r) = v \). This concludes the proof of the theorem.

REFERENCES


INTERNATIONAL BUSINESS MACHINES CORPORATION, YORKTOWN HEIGHTS, NEW YORK