Topological Entropy

In this lecture we define and give the basic properties of topological entropy. We start with the original definition due to Adler, Konheim and McAndrew. It has much in common with measure-theoretic entropy, but instead of partitions it uses open covers for compact metric space.

Definition: We say that $\mathcal{U} = \{U_i\}$ is an *open* ε*-cover* if all U_i are open sets of diameter $\leq \varepsilon$ and $X \subset \bigcup_i U_i$.

Given two open covers $\mathcal{U}$ and $\mathcal{V}$, the joint

$$\mathcal{U} \vee \mathcal{V} := \{U \cap V : U \in \mathcal{U}, V \in \mathcal{V}\}$$

is an open cover again.

Given a continuous map $T : X \to X$, the n-th joint of $\mathcal{U}$ is

$$\mathcal{U}^n := \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{U}).$$

NB: Without continuity of T, $T^{-1}(\mathcal{U})$ need not be an open cover.

Topological Entropy

A subcover of $\mathcal{U}$ is a subcollection of $\mathcal{U}$ that still covers X. Let

$$\mathcal{N}(\mathcal{U}) = \min\{\#\mathcal{V} : \mathcal{V} \text{ is subcover of } \mathcal{U}\}.$$

Note that by compactness of X, every open cover has a finite subcover, so $\mathcal{N}(\mathcal{U}) < \infty$.

Define the topological entropy as

$$h_{top}(T) = \lim_{\varepsilon \to 0} \sup_{\mathcal{U}} \lim_{n} \frac{1}{n} \log \mathcal{N}(\mathcal{U}^n), \tag{1}$$

where the supremum is taken over all open ε-covers $\mathcal{U}$.

Because $\mathcal{N}(\mathcal{U} \vee \mathcal{V}) \leq \mathcal{N}(U)\mathcal{N}(\mathcal{V})$, the sequence $(\log \mathcal{N}(\mathcal{U}^n))_{n \in \mathbb{N}}$ is subadditive, and $\lim_n \frac{1}{n} \log \mathcal{N}(\mathcal{U}^n)$ exists by Fekete's Lemma.

Topological Entropy

Let $T : X \to X$ be a continuous map on a compact metric space X.

Lemma 1

- $h_{top}(T^k) = kh_{top}(T)$ for $k \geq 0$.
- If T is invertible, then $h_{top}(T^{-1}) = h_{top}(T)$.

Topological Entropy

Proof: Let $\mathcal{U}$ be an open cover of X and $\mathcal{V} = \mathcal{U}^k$. Then $\mathcal{U}^{kn} = (\mathcal{V})^n$, and the exponential growth-rates:

$$h_{top}(T^k, \mathcal{V}) = k h_{top}(T, \mathcal{U}).$$

Because there might be open covers $\mathcal{V}$ that cannot be written as $\mathcal{V} = \mathcal{U}^k$, this only proves $h_{top}(T^k) \geq k h_{top}(T)$.

But $\mathcal{V} = \mathcal{U}^k$ refines $\mathcal{U}$, so that $h_{top}(T^k, \mathcal{U}) \leq h_{top}(T^k, \mathcal{V})$. Therefore also $h_{top}(T^k) \leq k h_{top}(T)$.

Exercise: Prove the second statement: If T is invertible, then $h_{top}(T^{-1}) = h_{top}(T)$.

Topological Entropy

Two maps (X, T) and (Y, S) are conjugate if there is a homeomorphism $h : X \to Y$ such that $h \circ T = S \circ h$. They are called semi-conjugate (and (Y, S) is a topological factor of (X, S)) if the map h is only continuous (and not necessarily with a continuous inverse).

Lemma 2 If (Y, S) is semi-conjugate to (X, T), then $h_{top}(S) \leq h_{top}(T)$. In particular, conjugate systems (on compact spaces!) have the same entropy.

Proof: Let $\mathcal{V}$ be an open cover of Y. Then $\mathcal{U} := h^{-1}(\mathcal{V})$ is an open over of X. Furthermore, indicating the map used as subscript:

$$h^{-1}(\mathcal{V}_S^n) = (\mathcal{U}_T)^n \text{ and } \mathcal{N}(\mathcal{V}^n) = \mathcal{N}(\mathcal{U}^n).$$

Therefore $h_{top}(T, \mathcal{U}) = h_{top}(S, \mathcal{V})$.

Because there are potentially open covers $\mathcal{U}$ on X that do not come from some open $\mathcal{V}$ of Y (for example, $\mathcal{V}' = \mathcal{V} \vee \{\text{fixed } \varepsilon\text{-cover}\}$), we have

$$h_{top}(T) \geq h_{top}(S).$$

If h is a conjugacy, then we can reverse the role of T and S and get the other inequality too.

Topological Entropy of Interval Maps

Let $T : [0,1] \to [0,1]$ be an interval map. A maximal piece on which T is monotone is called a lap or branch. Unimodal maps are maps with two laps. The lap-number is denoted as $\ell(T)$.

Figure: Unimodal maps: a quadratic map and a tent map.

The variation of T is defined as

$$\mathrm{Var}(T) = \sup_{0 \le x_0 < \dots\ x_N \le 1} \sum_{i=1}^{N} |T(x_i) - T(x_{i-1})|,$$

where the supremum runs over all finite collections of points in $[0,1]$.

Topological Entropy of Interval Maps

There are various shortcuts to compute the entropy of a continuous map $T : [0,1] \to [0,1]$. The following result is due to Misiurewicz & Szlenk:

Theorem Let $T : [0,1] \to [0,1]$ have finitely many laps. Then

$$\begin{aligned} h_{top}(T) &= \lim_{n\to\infty} \frac{1}{n} \log \ell(T^n) \\ &= \limsup_{n\to\infty} \frac{1}{n} \log \#\{\text{clusters of } n\text{-periodic points}\} \\ &= \max\{0, \lim_{n\to\infty} \frac{1}{n} \log \mathrm{Var}(T^n)\}. \end{aligned}$$

where two n-periodic points are in the same cluster if they belong to the same lap of T^n.

Topological Entropy of Interval Maps

Remark The identity map has one branch, consisting of (uncountably many) fixed point, that form one cluster. The map $x \mapsto x + (x/10)^2 \sin(\pi/x) \mod 1$ has also one branch, but with countably many fixed point, forming one cluster. For an expanding map, every branch can contain only one fixed point.

Remark From the variation part of the theorem, it follows immediatley that a continuous map with slope $\pm s$ (such as a tent map) has entropy $h_{top}(T) = \max\{\log s, 0\}$.

Topological Entropy of Interval Maps

Sketch of Proof: Since the variation of a monotone function is given by $\sup T - \inf T$, and due to the definition of "cluster" of n-periodic points, we have

$$\#\{\text{clusters of } n\text{-periodic points}\}, \mathrm{Var}(T^n) \leq \ell(T^n).$$

For a lap J of T^n, let $\gamma := |T^n(J)|$ be its *height*. We state without proof: For every $\delta > 0$, there is $\gamma > 0$ such that

$$\#\{J : J \text{ is a lap of } T^n, \ |T^n(J)| > \gamma\} \geq (1-\delta)^n \ell(T^n).$$

This means that $Var(T^n) \geq \gamma(1-\delta)^n \ell(T^n)$, and therefore

$$-2\delta + \lim_{n\to\infty} \frac{1}{n} \log \ell(T^n) \leq \lim_{n\to\infty} \frac{1}{n} \log \mathrm{Var}(T^n) \leq \lim_{n\to\infty} \frac{1}{n} \log \ell(T^n).$$

Since δ is arbitrary, $\lim_n \frac{1}{n} \log \mathrm{Var}(T^n) = \lim_n \frac{1}{n} \log \ell(T^n)$.

Topological Entropy of Interval Maps

Proof continued: Assume further[1] that there is $K = K(\gamma)$ such that $\cup_{i=0}^{K} T^i(J) = X$ for every lap of height $|T^n(J)| \geq \gamma$,

$$\#\{\text{clusters of } n+i\text{-periodic points}, 0 \leq i \leq K\} \geq (1-\delta)^n \ell(T^n).$$

This implies that

$$-2\delta \ + \lim_n \frac{1}{n} \log \ell(T^n) \leq$$
$$\limsup_{n\to\infty} \frac{1}{n} \max_{0\leq i\leq K} \log \#\{\text{clusters of } n+i\text{-periodic points}\},$$

so
$\lim_n \frac{1}{n} \log \ell(T^n) = \limsup_{n\to\infty} \frac{1}{n} \log \#\{\text{clusters of } n\text{-periodic points}\}$.

[1] Without proof. In fact, it is not entirely true if T has an invariant subset attracting an open neighbourhood. But it suffices to restrict T to its *nonwandering set*, that is, the set $\Omega(T) = \{x \in X : x \in \cup_{n\geq 1} T^n(U))$ for every neighbourhood $U \ni x\}$, because $h_{top}(T) = h_{top}(T|_{\Omega(T)})$.

Topological Entropy of Interval Maps

Proof continued: If $\varepsilon > 0$ is so small that the width of every lap is greater than 2ε, then for every ε-cover $\mathcal{U}$, every subcover of $\mathcal{U}^n$ has at least one element in each lap of T^n. Therefore $\ell(T^n) \leq \mathcal{N}(\mathcal{U}^n)$.

On the other hand, for this ε-cover $\mathcal{U}$, if N is so large that the width of every lap of T^N is smaller than every $U \in \mathcal{U}$, then we also have $\mathcal{N}(\mathcal{U}^n) \leq \ell(T^{n+N})$. Therefore

$$\lim_{n\to\infty} \frac{1}{n} \log \ell(T^n) \leq h_{top}(T) \leq \lim_n \frac{1}{n} \log \ell(T^{n+N}).$$

This shows that $h_{top}(T) = \lim_n \frac{1}{n} \log \ell(T^n)$.

Now all limits have been shown to be the same, so the proof sketch is complete.

Bowen's Approach

Let T be map of a compact metric space (X, d). If my eyesight is not so good, I cannot distinguish two points $x, y \in X$ if they are at a distance $d(x, y) < \varepsilon$ from one another. I may still be able to distinguish there orbits, if $d(T^k x, T^k y) > \varepsilon$ for some $k \geq 0$. Hence, if I'm willing to wait $n - 1$ iterations, I can distinguish x and y if

$$d_n(x, y) := \max\{d(T^k x, T^k y) : 0 \leq k < n\} > \varepsilon.$$

If this holds, then x and y are said to be (n, ε)-separated.

Among all the subsets of X of which all points are mutually (n, ε)-separated, choose one, say $E_n(\varepsilon)$, of maximal cardinality. Then $s_n(\varepsilon) := \#E_n(\varepsilon)$ is the maximal number of n-orbits I can distinguish with ε-poor eyesight.

Bowen's Approach

The topological entropy is defined as the limit (as $\varepsilon \to 0$) of the exponential growth-rate of $s_n(\varepsilon)$:

$$h_{top}(T) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s_n(\varepsilon). \tag{2}$$

Note that $s_n(\varepsilon_1) \geq s_n(\varepsilon_2)$ if $\varepsilon_1 \leq \varepsilon_2$, so $\limsup_n \frac{1}{n} \log s_n(\varepsilon)$ is a decreasing function in ε, and the limit as $\varepsilon \to 0$ indeed exists.

Bowen's Approach

Instead of (n, ε)-separated sets, we can also work with (n, ε)-spanning sets, that is, sets that contain, for every $x \in X$, a y such that $d_n(x, y) \leq \varepsilon$. Due to its maximality, $E_n(\varepsilon)$ is always (n, ε)-spanning, and no proper subset of $E_n(\varepsilon)$ is (n, ε)-spanning. Set

$$r_n(\varepsilon) = \min\{F_n(\varepsilon) : F_n(\varepsilon) \text{ is } (n, \varepsilon)\text{-spanning}\}$$

Each $y \in E_n(\varepsilon)$ must have a point of an $(n, \varepsilon/2)$-spanning set within an $\varepsilon/2$-ball (in d_n-metric) around it, and by the triangle inequality, this $\varepsilon/2$-ball is disjoint from $\varepsilon/2$-ball centred around all other points in $E_n(\varepsilon)$. Therefore,

$$r_n(\varepsilon) \leq s_n(\varepsilon) \leq r_n(\varepsilon/2). \tag{3}$$

Thus we can equally well define

$$h_{top}(T) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r_n(\varepsilon). \tag{4}$$

Bowen's Approach

Examples: Consider the β-transformation

$$T_\beta : [0,1) \to [0,1), \quad x \mapsto \beta x \pmod 1$$

for some $\beta > 1$.

Take $\varepsilon < 1/(2\beta^2)$, and $G_n = \{\frac{k}{\beta^n} \ : \ 0 \le k < \beta^n\}$. Then G_n is (n, ε)-separating, so $s_n(\varepsilon) \ge \beta^n$.

On the other hand, $G_n' = \{\frac{2k\varepsilon}{\beta^n} \ : \ 0 \le k < \beta^n/(2\varepsilon)\}$ is (n, ε)-spanning, so $r_n(\varepsilon) \le \beta^n/(2\varepsilon)$. Therefore

$$\log \beta = \limsup_{n\to\infty} \frac{1}{n} \log \beta^n \le h_{top}(T_\beta) \le \limsup_{n\to\infty} \frac{1}{n} \log \beta^n/(2\varepsilon) = \log \beta.$$

NB: the β-transformation is not continuous, but Bowen's definition of entropy works.

Bowen's Approach

Examples: Circle rotations, or in general isometries, T have zero topological entropy. Indeed, if $E(\varepsilon)$ is an ε-separated set (or ε-spanning set), it will also be (n,ε)-separated (or (n,ε)-spanning) for every $n \geq 1$. Hence $s_n(\varepsilon)$ and $r_n(\varepsilon)$ are bounded in n, and their exponential growth rates are equal to zero.

Let (X,σ) be the full shifts on N symbols. Let $\varepsilon > 0$ be arbitrary, and take m such that $2^{-m} < \varepsilon$. If we select a point from each $n+m$-cylinder, this gives an (n,ε)-spanning set, whereas selecting a point from each n-cylinder gives an (n,ε)-separated set. Therefore

$$\begin{aligned}
\log N = \limsup_{n\to\infty} \frac{1}{n} \log N^n \quad &\leq \quad \limsup_{n\to\infty} \frac{1}{n} \log s_n(\varepsilon) \\
&\leq \quad \limsup_{n\to\infty} \frac{1}{n} \log r_n(\varepsilon) \\
&\leq \quad \limsup_{n\to\infty} \log N^{n+m} = \log N.
\end{aligned}$$

Bowen's Approach

Proposition For a continuous map T on a compact metric space (X, d), the three definitions of topological entropy (1), (2) and (4) give the same outcome.

Proof: The equality of the limits (2) and (4) follows directly from (3):

$$r_n(\varepsilon) \leq s_n(\varepsilon) \leq r_n(\varepsilon/2).$$

If $\mathcal{U}$ is an ε-cover, every $A \in \mathcal{U}^n$ can contain at most one point in an (n, ε)-separated set, so $s(n, \varepsilon) < \mathcal{N}(\mathcal{U}^n)$, Therefore

$$\limsup_{n\to\infty} \frac{1}{n} \log s(n, \varepsilon) \leq \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(\mathcal{U}^n).$$

Bowen's Approach

Proof continued: In a compact metric space, every open cover $\mathcal{U}$ has a number (called its Lebesgue number) such that for every $x \in X$, there is $U \in \mathcal{U}$ such that $B_\delta(x) \subset U$. Clearly $\delta < \varepsilon$ if $\mathcal{U}$ is an ε-cover.

Now if an open ε-cover $\mathcal{U}$ has Lebesgue number δ, and E is an (n, δ)-spanning set of cardinality $\#E = r(n, \delta)$, then

$$X \subset \bigcup_{x \in E} \bigcap_{i=0}^{n-1} T^{-i}(B_\delta(T^i x)).$$

Since each $B_\delta(T^i(x))$ is contained in some $U \in \mathcal{U}$, we have $\mathcal{N}(\mathcal{U}^n) \leq r(n, \delta)$. Since $\delta \to 0$ as $\varepsilon \to 0$, also

$$\lim_{\varepsilon \to 0} \lim_n \frac{1}{n} \log \mathcal{N}(\mathcal{U}^n) \leq \lim_{\delta \to 0} \limsup_n \frac{1}{n} \log r(n, \delta).$$

This completes the proof.