The definition of topological entropy in terms of spanning numbers requires limsup: A simpler construction

Winfried Just
Department of Mathematics, Ohio University

Based on joint work with Ying Xin
John Hopkins University Medical School

Mathematical Biology and Dynamical Systems Seminar
Ohio University, September 10, 2019

Topological entropy of a dynamical system: The idea

For the purpose of this talk, a dynamical system is a pair (X, F), where X is a compact metric space with distance function D and $F : X \to X$ is a homeomorphism. Roughly speaking:

- A (forward) trajectory is a sequence $(F^t(x))_{t=0}^{\infty}$ for some $x \in X$.
- **In chaotic systems,** for sufficiently small $\varepsilon > 0$, the number $N_T(\varepsilon, D)$ of trajectories that are distinguishable at resolution ε within T time steps scales like $B(\varepsilon)^T$ for some $B(\varepsilon) > 1$.
- Thus we can use the growth rate of $N_T(\varepsilon, D)$ to define measure for how chaotic the system is:

  $h(X, F) = \lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln N_T(\varepsilon, D)}{T}$.

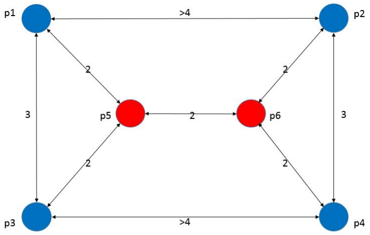
- This measure is called the topological entropy of (X, F).

Separation numbers and spanning numbers

Let (X, d) be metric space, and let $\varepsilon > 0$.

We define the separation number $sep(X, \varepsilon, d)$ as the **largest** size of a subset $Y \subset X$ such that $d(x, x') \geq \varepsilon$ for all $x, x' \in Y$,

and the spanning number $span(X, \varepsilon, d)$ as the **smallest** size of a subset $Y \subset X$ such that for all $x \in X$ there exists $y \in Y$ with $d(x, y) < \varepsilon$.



A = {p1, p2, p3, p4} is 3-separated of size 4 = sep(X, 3, d)
B = {p5, p6} is 3-spanning of size 2 = span(X, 3, d)

Two definitions of $N_T(\varepsilon, D)$

We could define $N_T(\varepsilon, D)$ as the largest size $sep(X, \varepsilon, D_T)$ of a (T, ε)-separated subset of X, that is, of a set $Y \subseteq X$ such that for all $x, x' \in Y$ there exists a $0 \le t < T$ with $D(F^t(x), F^t(x')) \ge \varepsilon$.

Or we could define $N_T(\varepsilon, D)$ as the smallest size $span(X, \varepsilon, D_T)$ of a (T, ε)-spanning subset of X, that is, of a set $Y \subseteq X$ such that for all $x \in X$ there exists $y \in Y$ such that for all $0 \le t < T$ we have $D(F^t(x), F^t(y)) < \varepsilon$.

The separation numbers $sep(X, \varepsilon, D_T)$ and spanning numbers $span(X, \varepsilon, D_T)$ are always finite and satisfy the inequality $sep(X, \varepsilon, D_T) \ge span(X, \varepsilon, D_T)$.

But why limsup?

So we can define the topological entropy $h(X, F)$ as:

$\lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln sep(X,\varepsilon,D_T)}{T} = \lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln span(X,\varepsilon,D_T)}{T}$.

But why $\limsup$? Could we use $\lim$ instead?

Chorus of the experts: It doesn't really matter. There is another definition, based on covering numbers, where you can.

But for the definitions based on separation or spanning numbers?

Chorus: Then you cannot.

But why?

Chorus: Because there is no obvious reason why you could.

Any counterexamples?

Chorus: Hmmm. Not that we know of.

(This was what we heard about 3 years ago.)

Our main theorem

So we went ahead and constructed one.

Theorem

There exists a system (X, F) with a metric D on X such that for $\varepsilon > 0$ we have:

$$\liminf_{T\to\infty} \frac{\ln sep(X, \varepsilon, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln sep(X, \varepsilon, D_T)}{T},$$

$$\liminf_{T\to\infty} \frac{\ln span(X, \varepsilon, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln span(X, \varepsilon, D_T)}{T}.$$

The system (X, F) is minimal; i.e., every forward trajectory is dense in X.

Remark: We have more results along these lines. They are included in Ying's dissertation. But the blue part was proved more recently.

A problem with this construction

Our proof of this therem is very long.

Even the definition of the space X and of the metric D on X take up several pages.

Our arXiv preprint of Ying's dissertation chapter has 69 pages. After some simplifications, a recently submitted journal version still has 54 pages.

Chorus of Potential Reviewers: (Show remarkable lack of enthusiasm for reading the proof.)

There must be a simpler example!

The authors: We certainly tried, but ...

What would a really, really simple example look like?

Let $X^+ = {}^{\mathbb{Z}}A$ be the space of two-sided sequences x of symbols from some finite alphabet A.

Consider the bijection $\# : \mathbb{Z} \to \mathbb{N}$ given by

$\#(0) = 0, \quad \#(1) = 1, \quad \#(-1) = 2, \quad \#(2) = 3, \quad \#(-2) = 4, \ldots$

and the function $\Delta : \left({}^{\mathbb{Z}}A\right)^2 \to \mathbb{N} \cup \{\infty\}$ that takes the value $\Delta(y, z) = \infty$ when $y = z$ and the value $\Delta(y, z) = \#(i)$ when $y \neq z$, where i is such that
$y(i) \neq z(i) \quad \text{and} \quad \forall j \in \mathbb{Z}\ (\#(j) < \#(i) \ \Rightarrow\ y(j) = z(j))$.

Define a metric ρ on ${}^{\mathbb{Z}}A$ as follows: $\rho(y, z) = \kappa^{-\Delta(y,z)}$, where $\kappa > 1$ and $\kappa^{-\infty} = 0$.

Let $\sigma : X^+ \to X^+$ be the shift operator defined by
$\sigma(x)(i) = x(i + 1)$ for all $i \in \mathbb{Z}$.

Then (X^+, ρ) is a full shift. It is a compact metric space and $\sigma : X^+ \to X^+$ is a homeomorphism.

Now let $X \subset {}^{\mathbb{Z}}A$ be closed both topologically wrt to ρ and wrt to σ. Then X is a subshift and (X, σ) is a standard subshift system.

Could we use a standard subshift system in our construction?

No. This is a well-known result.

But let d be any metric on A consider the following metric D on ${}^{\mathbb{Z}}A$:

$D(y, z) = d(y(\Delta(y, z)), z(\Delta(y, z)))\kappa^{-\Delta(y,z)}$,

where κ is sufficiently large so that this is actually a metric.

Then ρ and D generate the same topology on X^+, and when X is a subshift, we obtain a (nonstandard) subshift system (X, σ) with (nonstandard) subshift metric D. We will from now on call D the near-subshift metric and (X, σ, D) as above a near-subshift system.

Could we use a near-subshift system in our construction?

Some remarks on the terminology

We should really write $(X, \sigma \restriction X, D \restriction X^2)$, but we omit the extra symbols in the interest of reducing clutter.

While standard subshift systems have been extensively studied in the literature, the phrase "near-subshift" was coined by us. As far as we know, these systems had not been studied before. This is not surprising, as topologically they are the same as the standard ones.

The notion of a near-subshift metric is defined relative to an underlying metric d on A. The standard metric ρ is a special case for $d(a, b) = 1$ whenever $a \neq b$. In our construction we will use $A = \{0, 1, 2, 4, 5\}$ with $d(a, b) = \min\{|a - b|, 2\}$ and $\kappa = 3$.

We had shown some time ago that examples for any parts of our theorem with X being any nonstandard subshift system exist if, and only if, there exist corresponding examples with X being a subset of $\mathbb{R}^n$ with the metric induced by the sup-norm.

So, could we use a near-subshift system in our construction?

We actually started out trying to construct a near-subshift system (X, σ, D) with

$$\liminf_{T\to\infty} \frac{\ln sep(X, \varepsilon, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln sep(X, \varepsilon, D_T)}{T}.$$

This didn't work. We always seemed to run into the same technical issue that we still don't fully understand.

At this time we conjecture that there are no such examples, but the proof of our conjecture may require new techniques.

Only quite recently did we seriously try to construct a counterexample for the spanning numbers.

Near-subshift systems can be counterexamples for the spanning numbers

Theorem

There exists a near-subshift system (Z, σ, D) such that

$$\liminf_{T\to\infty} \frac{\ln span(Z, 2, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln span(Z, 2, D_T)}{T}.$$

Proof: We start by picking sequences $(T(n))_{n\in\mathbb{N}}$ and $(T^+(n))_{n\in\mathbb{N}}$ of positive integers such that the following inequalities hold:
$T(0) < T^+(0) < T(1) < T^+(1) < \cdots < T(n) < T^+(n) < \dots$
The idea is to construct the system in such a way that

$$\limsup_{n\to\infty} \frac{\ln span(Z, 2, D_{T(n)})}{T(n)} \leq 0.6 \ln 2,$$

$$\liminf_{n\to\infty} \frac{\ln span(Z, 2, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2.$$

Outline of our construction

Our alphabet A and the metric d on A here will be given by:

$$A = \{0, 1, 2, 4, 5\},$$
$$d(a, b) = \min\{|a - b|, 2\}.$$

We define D based on d with $\kappa = 3$.

We will construct Z as the union of two disjoint subshifts $Z = X \dot{\cup} Y$ such that:

- For every $n \in \mathbb{N}$, we can find $(T(n), 2)$-spanning sets for Z of size at most $2^{0.6T(n)}$ that heavily relies on elements of X,
- For every $n > 0$, every $(T^+(n), 2)$-spanning sets for Z must contain a subset of Y of size at least $2^{0.9T^+(n)}$.

Terminology: A few words about words

- A word of length T in A is a finite sequence $\psi \in {}^{T}A$.
- For a subset $I \subset \mathbb{Z}$ and an integer t we let $I + t = \{i + t : \ i \in I\}$.
- We extend the shift operator σ to the family of words in A, so that when ψ is a word of length T, then $\sigma^t(\psi)$ is a function defined on the set $T - t$ (not necessarily a word) that takes the values $\sigma^t(\psi)(j) = \psi(j + t)$ for all $j \in T - t$.
- Let φ, ψ be words of lengths $T^- \leq T$, respectively. Then φ is a subword of ψ if there exists t with $0 \leq t \leq T - T^-$ such that $\varphi = \sigma^t(\psi) \restriction T^-$. Similarly, φ is a subword of $z \in {}^{\mathbb{Z}}A$ if there exists some $t \in \mathbb{Z}$ such that $\varphi = \sigma^t(z) \restriction T^-$.
- Let φ, ψ be words of lengths T, T', respectively. Then the concatenation $\varphi^\frown\psi$ is the word of length $T + T'$ such that $\varphi^\frown\psi \restriction T = \varphi$ and $\sigma^T(\varphi^\frown\psi) \restriction T' = \psi$.

Background: A few more words about words

A standard argument shows that $X \subseteq {}^{\mathbb{Z}}A$ is a subshift if, and only if, $X \neq \emptyset$ and there exists a (possibly empty) set W^- of forbidden words such that X consists of all elements of ${}^{\mathbb{Z}}A$ that do not contain any element of W^- as a subword.

Conversely, if X is any subshift, then we can define, for each $T \geq 0$, the set $W_T^+ = W_T^+(X)$ of all permitted words of length T as the set of all words that are subwords of some $x \in X$.

The size of this set is closely related to the spanning and separation numbers: Let (X, σ, D) be any near-subshift system with $X \subseteq {}^{\mathbb{Z}}A$. Let D be defined based on any metric d on A with $\min\{d(a,b) : \ a \neq b \in A\} = 1$. Then

$$\forall T > 0 \quad sep(X, 1, D_T) = span(X, 1, D_T) = |W_T^+(X)|.$$

Permitted words in our construction

In order to define the subshift X, we will construct the families W_T^X and W_T^Y of permitted words of length T for the subshifts X and Y, respectively.

We will do this in stages, by first constructing families $\mathcal{X}_n^-, \mathcal{Y}_n^-$ of words of length $T(n)$ and $\mathcal{X}_n, \mathcal{Y}_n$ of words of length $T^+(n)$. Then ψ will be a permitted word of length $T(n)$ for X iff ψ is a subword of a concatenation of two words in $\mathcal{X}_n^-$, and ψ will be a permitted word of length $T^+(n)$ for X iff ψ is a subword of a concatenation of two words in $\mathcal{X}_n$. Analogously for Y.

We will need another parameter, a fixed sequence $(\alpha_n)_{n\in\mathbb{N}}$ of reals such that

- $0.1 = \alpha_0 \geq \alpha_n > 2\alpha_{n+1} > 0$ for all $n \in \mathbb{N}$.
- $\lim_{n\to\infty} \alpha_n = 0$.

Permitted words for X

- The family $\mathcal{X}_0^- := \{\psi_0, \varphi_0\} \subset {}^{T(0)}\{1,4\}$ consists of the function ψ_0 that is constantly equal to 1, and the function φ_0 that takes the value $\varphi_0(0) = 4$ and the value $\varphi_0(i) = 1$ for $0 < i < 10$.
- For any $n \in \mathbb{N}$, the set $\mathcal{X}_n$ consists of all words of length $T^+(n)$ that contain at most $\alpha_{n+1} T^+(n)$ occurrences of 4 and that are concatenations of successive words in $\mathcal{X}_n^-$.
- For any $n \in \mathbb{N}$, the set $\mathcal{X}_{n+1}^-$ consists of all words of length $T(n+1)$ that are concatenations of successive words in $\mathcal{X}_n$.

We will choose all numbers $T(n)$ so that they are multiples of 10. Then we get for every n:

$$span(X, 2, D_{T(n)}) \leq |W^X_{T(n)}| \leq 2^{0.1\,T(n)}.$$

Regular permitted words for Y

- The family $\mathcal{Y}_0^-$ consists of all functions $\psi \in {}^{T(0)}\{0,1,2,5\}$ such that $\psi(0) \in \{1,5\}$ and $\psi(i) \in \{0,2\}$ for $0 < i < 10$.
- We select one element $\psi_1 \in \mathcal{Y}_0^-$ with $\psi_1(0) = 1$ that will serve special purposes.
- All words in $\mathcal{Y}_0^-$ are considered regular words.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_n$ consists of all words of length $T^+(n)$ that are of two kinds, regular or special:
  - Regular words $\varphi \in \mathcal{Y}_n$ that are concatenations of regular words in $\mathcal{Y}_n^-$.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_{n+1}^-$ consists of all words of length $T(n+1)$ that are of two kinds, regular or special:
  - Regular words $\varphi \in \mathcal{Y}_{n+1}^-$ that are concatenations of regular words in $\mathcal{Y}_n$ and satisfy

$$\exists \psi \in \mathcal{X}_{n+1}^- \quad \{i \in T(n+1) : \varphi(i) = 5\} = \{i \in T(n+1) : \psi(i) = 4\}.$$

"Almost" $(T(n), 2)$-spanning sets of Z

Let us consider a regular word $\varphi \in \mathcal{Y}_n^-$. Then there exists a word $\psi \in \mathcal{X}_n^-$ such that $\psi(i) = 1$ whenever $\varphi(i) \in \{0, 2\}$, and $\psi(i) = 4$ whenever $\varphi(i) = 5$. Thus $D_{T(n)}(\varphi, \psi) = 1$.

In an ideal world where $W^Y_{T(n)}$ would consist only of regular words in $\mathcal{Y}_n^-$, this would imply that we could choose a $(T(n), 2)$-spanning set for X that consists entirely of elements of X, so that $span(Z, 2, D_{T(n)}) \leq 2^{0.1T(n)}$.

In the real world this doesn't quite work, because:

- We need to also deal with words φ that are concatenations of two words in $\mathcal{Y}_n^-$. This is straightforward but a bit technical and requires some additional conditions on the numbers $T(n)$ and $T^+(n)$. We will omit details here.
- Our definition of the subshift Y in terms of permitted rather than forbidden words requires us to also include some special words so that we do get a subshift.

Special permitted words for Y

We choose our parameters in such a way that all ratios $C(n) := T^+(n)/T(n)$ and $K(n+1) := T(n+1)/T^+(n)$ are odd integers.

- The family $\mathcal{Y}_0^-$ consists of all functions $\psi \in {}^{T(0)}\{0,1,2,5\}$ such that $\psi(0) \in \{1,5\}$ and $\psi(i) \in \{0,2\}$ for $0 < i < 10$.
- We select one element $\psi_1 \in \mathcal{Y}_0^-$ with $\psi_1(0) = 1$ that will serve special purposes.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_n$ consists of all words of length $T^+(n)$ that are of two kinds, regular or special:
  - Special words of the form $\varphi = \eta^\frown\psi^\frown\eta$, where $\psi \in \mathcal{Y}_n^-$ and η is a concatenation of $(C(n)-1)T(n)/20$ copies of ψ_1.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_{n+1}^-$ consists of all words of length $T(n+1)$ that are of two kinds, regular or special:
  - Special words of the form $\varphi = \eta^\frown\psi^\frown\eta$, where $\psi \in \mathcal{Y}_n$ and η is a concatenation of $(K(n+1)-1)T^+(n)/20$ copies of ψ_1.

Irregular words for Y

A word can be both regular and special in the sense of our definition. Special words that are not regular will be called irregular.

Irregular words are a bit like irregular verbs in learning a natural language: A nuisance to deal with, but fortunately there aren't too many of them.

Let $V^Y_{T(n)}$ denote the set of irregular words in $W^Y_{T(n)}$. We got the following estimate:

$$|V^Y_{T(n)}| \leq 10T(n)2^{(T(n)+T^+(n-1))/2}.$$

Under suitable technical assumptions on $T(n)$ and $T^+(n-1)$, this allows us to include all irregular words, together with all words in $W^X_{T(n)}$ in any spanning set without exceeding the upper bound $2^{0.6T(n)}$ on its size.

How about $span(Z, 2, D_{T^+(n)})$?

Recall some of our definitions:

- For any $n \in \mathbb{N}$, the set $\mathcal{X}_n$ consists of all words of length $T^+(n)$ that contain at most $\alpha_{n+1} T^+(n)$ occurrences of 4 and that are concatenations of successive words in $\mathcal{X}_n^-$.
- Regular words $\varphi \in \mathcal{Y}_n$ are concatenations of regular words in $\mathcal{Y}_n^-$.

This puts a restriction on the number of occurrences of 5 in regular words in $\mathcal{Y}_n$: By recursion and under some technical assumptions that are omitted here, we can show an upper bound of roughly $\alpha_n T^+(n)$ on their number.

But the number of occurrences of 4 in words in $W^X_{T^+(n)}$ is significantly smaller, there can be at most roughly $\alpha_{n+1} T^+(n)$ of them.

Thus $D_{T^+(n)}(x, y) = 2$ whenever $x \in X$ and $y \in Y$, so that $span(Z, 2, D_{T^+(n)}) \geq span(Y, 2, D_{T^+(n)})$.

How about $span(Y, 2, D_{T^+(n)})$?

Consider $y, y' \in Y$ such that $D_{T^+(n)}(y, y') \leq 1$ and y, y' take the value 5 for at least one $0 \leq i < T^+(n)$ each.

Let $S(y, n) := \{i < T^+(n) : y(i) \in \{0, 1]\} = S(y', n)$ and $S(y', n) := \{i < T^+(n) : y'(i) \in \{0, 1]\}$.

Then $S(y, n) = S(y', n)$.

Moreover, $y \restriction S(y, n) = y' \restriction S(y, n)$.

Also note that there are at least $2^{0.9T^+(n)}$ choices for $y'' \in Y$ with $S(y, n) = S(y'', n)$ such that y'' takes the value 5 for at least one $0 \leq i < T^+(n)$.

It follows that $span(Z, 2, D_{T^+(n)}), span(Y, 2, D_{T^+(n)}) \geq 2^{0.9T^+(n)}$.

This completes the proof of our theorem. $\square$