On the role of limsup in the definition of topological entropy: An alternative view of the construction

Winfried Just
Department of Mathematics, Ohio University

Based on joint work with Ying Xin
Montana State University

Mathematical Biology and Dynamical Systems Seminar
Ohio University, February 5, 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:

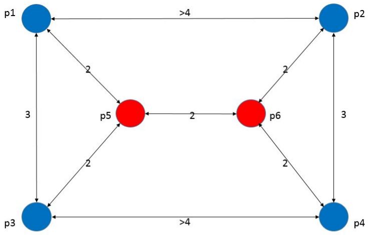
- A (forward) trajectory is a sequence $(F^t(x))_{t=0}^{\infty}$ for some $x \in X$.
- **In chaotic systems,** for sufficiently small $\varepsilon > 0$, the number $N_T(\varepsilon, D)$ of trajectories that are distinguishable at resolution ε within T time steps scales like $B(\varepsilon)^T$ for some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

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 of $Y \subset X$ 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 of $Y \subset X$ 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 $0 \leq t < T$ with $D(F^t(x), F^t(x')) \geq \varepsilon$.

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 $0 \leq t < T$ 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 $sep(X, \varepsilon, D_T) \geq span(X, \varepsilon, D_T)$.

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 lim sup? 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 is very new.

What did we do with these results?

1. Included as a chapter in Ying's dissertation.
2. Published all the details as a preprint:
   W. Just and Y. Xin (2017). On the role of limsup in the definition of topological entropy via spanning or separation numbers. Part I: Basic examples. *Preprint.* arXiv:1707.09052 (69 pages)
3. Submitted to journals.

Chorus of Potential Reviewers:

(Remarkable lack of enthusiasm for reading the proof.)

How can we make the presentation of this construction more palatable?

A normal way to present such a construction

1. Define the space X.
2. Define the function F.
3. Define the metric D on X.
4. Prove that these objects have all the desired properties.

Problem with this approach: We tried that one and tended to get a lot of discussion and feedback on steps 1 and 2.

But with a few notable exceptions, our audience didn't make it through step 3.

An alternative way of presenting our construction

Here is our proposed remedy:

1. Don't define X and D until half way into the proof.
2. Start with some relatively easy properties that our system (X, F) and the metric D will have, and then prove the inequalities (1) and (2) of the theorem from these properties.
3. Then formally define (X, F) and D with the help of certain parameters and show that they also satisfy the more pedestrian requirements of the theorem, like compactness of (X, D).
4. Finally show that suitable parameters for the definition of X and D actually exist.

The first ingredient of our construction

We will need positive integers $T(n)$ and $T^+(n)$ with

$1 < T(0) < T^+(0) < \cdots < T(n) < T^+(n) < T(n+1) < T^+(n+1) < \dots$

These sequences will be parameters in our construction and will satisfy a number of technical conditions that mostly boil down to requiring that $T(n+1)$ is sufficiently large relative to $T^+(n)$, and $T^+(n)$ is sufficiently large, but not too large, relative to $T(n)$.

We will then prove, for a suitably chosen parameter $\varepsilon > 0$, the following inequalities:

$$\liminf_{n\to\infty} \frac{\ln sep(X, \varepsilon, D_{T^+(n)})}{T^+(n)} > \limsup_{n\to\infty} \frac{\ln sep(X, \varepsilon, D_{2T(n)})}{2T(n)}.$$

$$\liminf_{n\to\infty} \frac{\ln span(X, \varepsilon, D_{T^+(n)})}{T^+(n)} > \limsup_{n\to\infty} \frac{\ln span(X, \varepsilon, D_{2T(n)})}{2T(n)}.$$

More precisely ...

More precisely, we will construct things so that for some $\lambda < 0.9$:

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

$$\limsup_{n\to\infty} \frac{\ln sep(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \lambda \ln 2.$$

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

$$\limsup_{n\to\infty} \frac{\ln span(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \lambda \ln 2.$$

What do we need to know about X and F?

In order to achieve this, we will construct (X, F) so that with every $x \in X$ and $n \in \mathbb{N}$ we can associate an integer $k_n(x) \in T^+(n) := \{0, \dots, T^+(n) - 1\}$ and a function $\psi_n(x) \in {}^{T^+(n)}\{0,1\}$ such that for all $x \in X$ and $n \in \mathbb{N}$:

(Fk) $k_n(F(x)) = k_n(x) + 1 \mod T^+(n)$

(Fx) If $k_n(x) < T^+(n) - 1$, then $\psi_n(x) = \psi_n(F(x))$.

We will require that $\psi_n \in \mathcal{X}_n$, where $\mathcal{X}_n$ is a subset of ${}^{T^+(n)}\{0,1\}$ that satisfies certain conditions, in particular, is of size at least $|\mathcal{X}_n| \geq 2^{0.9T^+(n)}$.

The key property of the metric D

The metric D on X is then defined in terms of a sequence of conditions $Cond_n$ on triplets (φ, ψ, k) such that

(Dε) $D(x, x') \geq \varepsilon$

if, and only if,

$$\forall n \in \mathbb{N} \;\; (k_n(x) \neq k_n(x') \;\vee\; Cond_n(\psi_n(x), \psi_n(x'), k_n(x))).$$

Keeping some separation numbers small

Next we want to assure that for some fixed $\lambda < 0.9$

$$\exists N \in \mathbb{N} \forall n > N \; span(X, \varepsilon, D_{2T(n)}) \leq sep(X, \varepsilon, D_{2T(n)}) < 2^{\lambda 2T(n)},$$

$$\exists N \forall n > N \; \frac{\ln span(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \frac{\ln sep(X, \varepsilon, D_{2T(n)})}{2T(n)} < \lambda \ln 2.$$

As long as $T^+(n) \leq T(n)2^{0.01T(n)}$ it suffices to show that

$$sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)},$$

since for all $\lambda > 0.88$ and all sufficiently large n we will then have

$$sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)} \leq T(n)2^{1.76T(n)} < 2^{\lambda 2T(n)}.$$

In order to achieve this goal, we need to design the conditions $Cond_n$ in the definition of the metric D in a certain way that is based on so-called colorings.

Colorings

Let $C(n) = T^+(n)/T(n)$. We partition the interval $T^+(n) := \{0, \ldots, T^+(n) - 1\}$ into consecutive subintervals I_j^n of length $T(n)$ each, where j ranges from 1 to $C(n)$.

Let ${}^{T^+(n)}\{0,1\}$ denote the set of all functions with domain $T^+(n)$ that take values in the set $\{0,1\}$. For a subset $S \subseteq {}^{T^+(n)}\{0,1\}$, let $[S]^2$ denote the set of all unordered pairs $\{\varphi, \psi\}$ of different functions from S. Moreover, let $[C(n)] = \{1, 2, \ldots, C(n)\}$. For our purposes, a coloring will be a function $c : [S]^2 \to [C(n)]$, where $S \subseteq {}^{T^+(n)}\{0,1\}$ for some $n \in \mathbb{N}$.

Definition

A subset $S^- \subseteq S$ is ≤ 2-*chromatic* for c if the restriction of c to $[S^-]^2$ takes at most 2 values from the set $[C(n)]$.

The conditions $Cond_n$

In our construction we use a sequence $(c_n)_{n \in \mathbb{N}}$ of suitable colorings as parameters. In particular, the domain of c_n will be $[\mathcal{X}_n]^2$, and the colorings c_n will not admit large ≤ 2-chromatic subsets of size $\geq 2^{0.75T(n)}$.

We prove existence of suitable colorings using the probabilistic method.

The conditions $Cond_n$ for the definition of the metric D will then take the form:

$$Cond_n(\varphi, \psi, k_n) \Leftrightarrow (\varphi(k_n) \neq \psi(k_n) \ \& \ k_n \in I^n_{c_n(\varphi,\psi)}).$$

Let us next illustrate how this works for keeping some separation numbers small.

Consider a $(2T, \varepsilon)$-separated subset $A \subset X$ such that $k_n(x) = \tau$ for all $x \in A$. Since there are only $T^+(n)$ possible choices for τ, for the inequality $sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)}$ it will suffice to show that such A has size at most $2^{1.75T(n)}$.

Let us focus in this illustration on the simplest case where τ with $0 \leq \tau < \tau + 2T(n) < T^+(n)$.

Then by (Fk) and (Fx), for all $t \in 2T$ and $x \in A$ we have $k_n(F^t(x)) = \tau + t$ and $\psi_n(F^t(x)) = \psi_n(x)$.

Since A was assumed to be $(2T, \varepsilon)$-separated, for every $x \neq x' \in A$ there exists $t < 2T$ such that $D(F^t(x), F^t(x')) \geq \varepsilon$.

By (Dε), this implies that for every $x \neq x' \in A$ there exists $t < 2T$ such that $Cond_n(\psi_n(F^t(x)), \psi_n(F^t(x')), \tau + t)$ holds. In view of (16) this now implies that for every $x \neq x' \in A$ there exists $t < 2T$ such that $\psi_n(x)(t + \tau) \neq \psi_n(x')(t + \tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x), \psi_n(x'))}$.

Small separation numbers, continued

Now assume towards a contradiction that A has more than $2^{1.75T(n)}$ elements and for every $x \neq x' \in A$ there exists $t < 2T$ such that $\psi_n(x)(t+\tau) \neq \psi_n(x')(t+\tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

Then there exists a subset $A^- \subset A$ of size $> 2^{0.75T(n)}$ such that $\psi_n(x), \psi_n(x')$ take the same values on $\tau + t$ for all $\tau + T(n) \leq t < \tau + 2T(n)$, so that for $x \neq x' \in A^-$ there exists $t < T$ such that $\psi_n(x)(t+\tau) \neq \psi_n(x')(t+\tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

But note that since each of the intervals I^n_j has length T, the values $\tau, \tau+1, \dots, \tau+T-1$ can belong to at most two of the intervals I^n_j.

It follows that the set $S := \{\psi_n(x) : \ x \in A^-\}$ has the same size as the set A^- and that the restriction of c_n to the set $[S]^2$ takes at most 2 values from the set $C(n)$. Thus the size of A^- cannot exceed the maximal size of a ≤ 2-chromatic subset of c_n, which contradicts our choices of A^- and c_n.

Large $(T^+(n), \varepsilon)$-separated subsets of X

For each n, we will construct a set W_n such that:

$\forall \psi \in \mathcal{X}_n \, \exists x \in W_n \;\; \psi_n(x) = \psi$. Thus $|W_n| \geq 2^{0.9T^+(n)}$.

$\forall m \in \mathbb{N} \, \forall x \in W_n \;\; k_m(x) = 0$.

$\forall x \neq x' \in W_n \, \exists t \in T^+(n) \, \forall m \in \mathbb{N} \; Cond_m(\psi_m(F^t(x)), \psi_m(F^t(x)), k_m))$,
where $k_m = t \mod T^+(m)$.

By Property (Fk), for $x, x' \in W_n$ we will then have
$k_m(F^t(x)) = k_m(F^t(x')) = t \mod T^+(m)$ for all m,
and Property (Dε) will imply that

$$sep(X, \varepsilon, D_{T^+(n)}) \geq 2^{0.9T^+(n)}, \text{ so that } \frac{\ln sep(X, \varepsilon, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2.$$

We will get the analogue of the last inequaity for spanning numbers by showing that for any $x' \in X$ the inequality $D_{T^+(n)}(x', x) < \varepsilon$ can hold for at most one $x \in W_n$.

How do we get large spanning numbers?

We will not define the sets W_n here. Instead, as we go, we will list additional properties that we need to get large spanning numbers.

Fix $n \in \mathbb{N}$ and let $x' \in X$. We will need the following property:

(X1) $x \neq x' \in W_n \quad \Longrightarrow \quad \psi_n(x) \neq \psi_n(x')$.

Then there exists at most one $x \in W_n$ with $\psi_n(x) = \psi_n(x')$.

So assume $x \in W_n$ is such that $\psi_n(x) \neq \psi_n(x')$. It suffices to show that $D_{T^+(n)}(x, x') \geq \varepsilon$.

For that we need some $t \in T^+(n)$ with $D(F^t(x), F^t(x')) \geq \varepsilon$.

Note that we need to find such a t that works at all levels m simultaneously.

Note also that the first clause of (Dε) essentially tells us that levels m with $k_m(x) \neq k_m(x')$ are unproblematic, so we will restrict our illustration here to the most interesting situation where $k_m(x') = k_m(x) = 0$ for all m. Then we also have $k_m(F^t(x')) = k_m(F^t(x))$ for all m and all t by Property (Fk).

How do we get large spanning numbers? Continued.

Now we need some t with $t = k_n(F^t(x)) = k_n(F^t(x'))$ such that

- $\psi_n(x)(t) \neq \psi_n(x')(t)$.
- $t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

We get such t from the following property of our colorings:

(CD) $\psi \neq \varphi \in \mathcal{X}_n$, then there exists at least one $k \in I^n_{c_n(\psi,\varphi)}$ such that $\psi(k) \neq \varphi(k)$.

Note then any $t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$ with $\psi_n(x)(t) \neq \psi_n(x')(t)$ will work for satisfying the clause in (Dε) that deals with level n.

We still need to find a t, in this same interval, that covers the clauses of (Dε) that deal with levels $m < n$ and with levels $m > n$.

Dealing with levels $m < n$

We need the following properties:

(TC) $T^+(n)$ is an integer multiple of $T^+(m)$ for $m < n$.

(XC) For $x \in X$, $m < n$, and $\ell T^+(m) < T^+(n)$, the restriction of $\psi_n(x)$ to $\{\ell T^+(m), \dots, (\ell+1)T^+(m) - 1\}$ is equal to $\psi_m(F^{\ell T^+(m)}(x))$.

Assume by induction that

- $\psi_n(x) \restriction I^n_{c_n(\psi_n(x),\psi_n(x'))} \neq \psi_n(x') \restriction I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

Let $m = n - 1$. Then we find $\tau = \ell T^+(m)$ with

- $\psi_m(F^\tau(x)) \neq \psi_m(F^\tau(x'))$.
- $k_m(x) = k_m(x') = 0$.

Now we can find $t' \in I^m_{c_m(\psi_m(F^\tau(x)),\psi_m(F^\tau(x')))}$ as on the previous slide, and $t = (c_n(\psi_n(x), \psi_n(x')) - 1)C(n) + \ell T^+(m) + t'$ will work at both levels m and n.

By iterating this argument, we find t that works at al levels $m \leq n$.

Dealing with levels $m > n$

The t we have found so far is an element of I_1^m for $m > n$.

This t will automatically work at levels $m > n$ if we have the following properties:

(WS) For $x \in W_m$ and $m < n$, the function $\psi_n(x)$ is "special."

(C1) If $\psi \neq \varphi \in \mathcal{X}_n$ and at least one of ψ, φ is special, then $c_n(\psi, \varphi) = 1$.

So what about that definition of (X, F)?

We let X consist of pairs $x = (y, \kappa)$, where $y \in {}^{\mathbb{Z}}\{0, 1\}$ and $\kappa : \mathbb{N} \to \mathbb{N}$ is a function such that $\kappa(n) \in T^+(n)$ for all n.

X will not consist of all such pairs, only of the pairs that are allowed by our conditions.

$F(y, \kappa) = (\sigma(y), \kappa \oplus 1)$, where:

- σ is the usual shift operator,
- $(\kappa \oplus 1)(n) = \kappa(n) + 1 \mod T^+(n)$ for all $n \in \mathbb{N}$.

We let $k_n(x) = \kappa(n)$ and $\psi_n(x) = \sigma^{-\kappa(n)}(y) \restriction T^+(n)$.

We then define the sets W_n so that they satisfy all relevant conditions.

And what about that definition of D?

We define D so that:

- (Dε) holds, and
- For each $x = (y, \kappa) \in X$ and $N \in \mathbb{N}$ the sets $U_N(x)$ of all $x' = (y', \kappa')$ such that $\forall i \; |i| < N \implies y(i) = y'(i) \;\&\; \kappa(i) = \kappa'(i)$ are open and form a basis for the topology.

The proof that X is compact and F a homeomorphism then becomes completely standard and independent from the more technical parts of the argument.

Minimality of (X, F) will result from the following properties:

(FC) When $x \in X$ and $m < n$, then $k_m(x) = k_n(x) \mod T^+(m)$.

(XM) For all $m < n$ and all $\psi, \varphi \in \mathcal{X}_m$, every $\psi^+ \in \mathcal{X}_n$ contains a block of the form $\psi^\frown\varphi$.