Będlewo, July 3, 2008

SYMBOLIC EXTENSIONS OF SMOOTH INTERVAL MAPS

(*THE ANTARCTIC THEOREM*)

Tomasz Downarowicz
(joint work with Alejandro Maass)

- - - D A Y   1 - - -

1. GENERALLY ON SYMBOLIC EXTENSIONS

Let (X, T) be a topological dynamical system, i.e., a continuous map on a compact metric space.

QUESTION 1: Is (X, T) a factor of a subshift?
In other words, does there exist a symbolic extension of (X, T)?

QUESTION 2: If yes, what is the infimum entropy of such a subshift?
Is this infimum attained?

A symbolic extension can be thought of as a lossless "digitalization" of the system.
In ergodic theory finite entropy is the only restriction to get an isomorphic symbolic representation (Krieger's Generator Theorem).
In topological dynamics this problem is much more complicated.
Finite partition coding usually destroys the topology and leads to loss of information.
Symbolic extension is the only solution that preserves everything (but adds "unwanted" dynamics).
Existence of symbolic (equivalently of an expansive) extension depends on subtle entropy properties.

Define

$$\mathbf{h}_{sex}(X,T) = \inf\{\mathbf{h}_{top}(Y,S) : (Y,S) \text{ is a two-sided subshift extension of } (X,T)\}$$
$$\mathbf{h}_{sex}(X,T) = \infty \text{ if } (X,T) \text{ has no symbolic extensions}$$

Let $\mathcal{P}_T(X)$ denote the set of all T-invariant measures μ on X.
Let $\pi : (Y,S) \to (X,T)$ be a factor map.
Define

$$h^\pi(\mu) = \sup\{h_\nu(S) : \nu \in \mathcal{P}_S(Y), \pi^*(\nu) = \mu\}$$

Define

$$h_{sex}(\mu) = \inf\{h^\pi(\mu) : \pi \text{ is a two-sided subshift extension of } (X,T)\}$$
$$h_{sex} \equiv \infty \text{ if } (X,T) \text{ has no symbolic extensions}$$

- -

PROBLEM 1.
Compute (or estimate) $\mathbf{h}_{sex}$ and h_{sex} for a given system (X,T) using its internal properties.

PROBLEM 2.
Estimate $\mathbf{h}_{sex}$ and h_{sex} for smooth maps on manifolds (C^r, $1 \le r \le \infty$).

2. HISTORY

1. Every expansive T has a symbolic extension (W. Reddy, 1968)

2. Not every finite entropy homeomorphism admits a symbolic extension
(M. Boyle ∼1990, published with D. Fiebig and U. Fiebig, 2002)

3. Formula for $\mathbf{h}_{sex}$ in dimension zero (T.D., 2001), in particular asymptotically h-expansive zero entropy systems have symbolic extensions with the same topological entropy.

4. Asymptotically h-expansive systems have symbolic extensions with the same entropy for every measure, i.e., $h_{sex} \equiv h$ (principal symbolic extension) (Boyle-Fiebig-Fiebig 2002)
This applies to C^∞ maps on manifolds (J. Buzzi, 1997).

5. Complete general description of h_{sex} in terms of "superenvelopes" and the "transfinite sequence"
(M. Boyle, T.D., 2004)
in particular:
Sex entropy variational principle

$$\mathbf{h}_{sex}(X,T) = \sup\{h_{sex}(\mu) : \mu \in \mathcal{P}_T(X)\}.$$

And also:
Attainment criterion (existence of a symbolic extension π realizing $h^\pi \equiv h_{sex}$):
if and only h_{sex} is affine.

6. Theory of entropy structures (T.D., 2005)
(we will refer to it later)

State of art for smooth systems

1. C^∞ implies asymptotic h-expansiveness which is equivalent to the condition $h_{sex} \equiv h$ (implying $\mathbf{h}_{sex} = \mathbf{h}_{top}$).

Successful application of entropy structures to smooth systems in dimension ≥ 2 (T.D., S. Newhouse, 2005):

2. C^1 admits systems with no symbolic extensions ($\mathbf{h}_{sex} = \infty$) – such systems are typical among area preserving non-Anosov diffeomorphisms.

3. C^r $(1 < r < \infty)$ admits systems where $\mathbf{h}_{sex} > \mathbf{h}_{top}$, typically $\mathbf{h}_{sex} \geq \mathbf{h}_{top} + \frac{R(f)}{r-1}$ $(R(f) = \limsup \frac{1}{n} \log \|Df^n\|)$

CONJECTURE: $\mathbf{h}_{sex} \leq \mathbf{h}_{top} + \frac{R(f)}{r-1}$.
In particular, every C^r map with $r > 1$ has a symbolic extension (is a factor of a subshift).

4. (David Burguet, 2008): C^r examples on the interval with $\mathbf{h}_{sex} \geq \mathbf{h}_{top} + \frac{R(f)}{r-1}$.

3. LATEST PROGRESS

(T.D., A. Maass) The conjecture holds in dimension 1 (interval, circle).

More precisely, the following estimate holds

Theorem
If X denotes the interval or the circle and $f : X \to X$ is a C^r map $(1 \le r \le \infty)$, then for every $\mu \in \mathcal{P}_f(X)$,

$$h_{sex}(\mu) \le h_\mu(T) + \frac{\bar{\chi}_0(\mu)}{r-1}$$

($\chi_0(\mu)$ for ergodic μ equals the maximum of 0 and the Lyapunov exponent; for other measures $\bar{\chi}_0$ is the average of χ_0 over the ergodic decomposition).

In particular,
every $C^{1+\epsilon}$ interval or circle map is a factor of a subshift.

4. INTRODUCTION TO ENTROPY STRUCTURES

I will NOT give the general definition of the entropy structure.

There are, however, many "particular" entropy structures.
Each is a sequence of functions $h_k : \mathcal{P}_T(X) \to [0, \infty)$, such that $h_k \nearrow h$ pointwise.
Sometimes it is better to consider the *tails* $\theta_k = h - h_k$. We have $\theta_k \searrow 0$ pointwise.

The derivation of h_{sex} from the entropy structure is via the "transfinite sequence".

Step 0: $u_0 \equiv 0$

Step $\alpha + 1$: $u_{\alpha+1} = \lim_k \widetilde{u_\alpha + \theta_k}$ (recall that $\tilde{f}(x) = \limsup_{y \to x} f(y)$)

Step β (limit ordinal): $u_\beta = \widetilde{\sup_{\alpha<\beta}} u_\alpha$

Theorem
There exists a countable ordinal α_0 such that $u_\alpha = u_{\alpha_0}$ for every $\alpha \geq \alpha_0$, and

$$h_{sex} = h + u_{\alpha_0}, \quad \mathbf{h}_{sex} = \sup_{\mathcal{P}_T(X)} (h + u_{\alpha_0}).$$

By the way, the famous Misiurewicz parameter $\mathbf{h}^*$ equals the pointwise supremum of the function u_1.

The Newhouse entropy structure

Definition. (Newhouse, 1989)

(a) $H(n,\delta|x,F,\mathcal{V}) := \log\max\{\#E : E$ is an (n,δ)-separated set in $F\cap V_x^n\}$;

(b) $H(n,\delta|F,\mathcal{V}) := \sup_{x\in F} H(n,\delta|x,F,\mathcal{V})$;

(c) $h(\delta|F,\mathcal{V}) := \limsup_n \frac{1}{n}H(n,\delta|F,\mathcal{V})$;

(d) $h(X|F,\mathcal{V}) := \lim_{\delta\to 0} h(\delta|F,\mathcal{V})$;

(e) for an ergodic measure ν, $h^{New}(X|\nu,\mathcal{V}) := \lim_{\sigma\to 1}\inf\{h(X|F,\mathcal{V}) : \nu(F) > \sigma\}$.

We extend the function $h^{New}(X|\cdot,\mathcal{V})$ to all of $\mathcal{P}_T(X)$ by averaging over the ergodic decomposition. This function is called the *local entropy function given the cover* $\mathcal{V}$.

The Newhouse entropy structure is obtained as the sequence

$$\theta_k(\mu) = h^{New}(X|\mu,\mathcal{V}_k),$$

where $\mathcal{V}_k$ is a sequence of open covers, each finer than the preceding one, and with the maximal diameters of their elements decreasing to zero. This is indeed an entropy structure (T.D. 2005).

5. KEY INGREDIENT IN THE ONE-DIMENSIONAL RESULT:

The Antarctic Theorem

Let f be a C^r transformation of the interval or of the circle X, where $r > 1$.
Let $\mu \in \mathcal{P}_f(X)$ and fix some $\gamma > 0$.

Then there exists an open cover $\mathcal{V}$ of X and a neighborhood of μ in $\mathcal{P}_f(X)$ such that for every ergodic ν in this neighborhood,

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{\chi_0(\mu) - \chi_0(\nu)}{r-1} + \gamma,$$

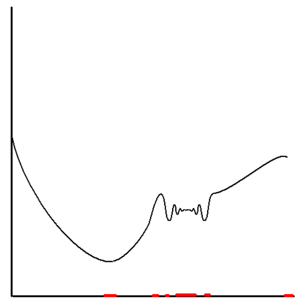
where

$$\chi(\mu) = \int \log |f'(x)|\, d\mu, \quad \chi_0(\mu) = \max\{0, \chi(\mu)\}$$

How one proves such a thing?

One needs to cleverly choose the cover $\mathcal{V}$ and the set F of measure ν close to 1:

$\mathcal{V}$ consists of one open set containing all critical points and finitely many intervals on which f is monotone.



F is the set where for n large enough the Cesaro means of the function $\log|f'|$ are close to $\chi(\nu)$.

Then the key calculation is in the following lemma:

Lemma
There exists a constant c such that for every $s > 0$, the number of monotone branches where $|f'|$ exceeds s is at most

$$c \cdot s^{-\frac{1}{r-1}}$$

6. Deducing the main result from the Antarctic Theorem

First, using a lot of functional analysis, we get rid of the assumption that ν is ergodic replacing the function χ_0 by $\bar{\chi}_0$ (this is **The Passage Theorem**): for ν near μ,

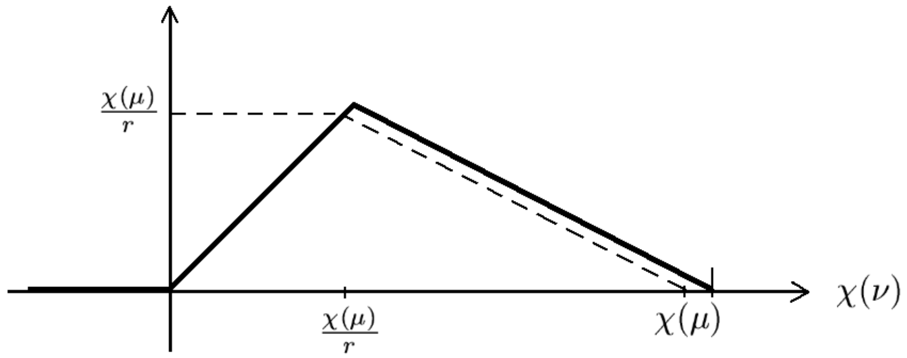
$$h^{New}(X|\nu, \mathcal{V}) \le \frac{\bar{\chi}_0(\mu) - \bar{\chi}_0(\nu)}{r-1} + \gamma.$$

Observe that also

$$h^{New}(X|\nu, \mathcal{V}) \le h(\nu) \le \bar{\chi}_0(\nu),$$

hence

$$h^{New}(X|\nu, \mathcal{V}) \le \min\{\bar{\chi}_0(\nu), \frac{\bar{\chi}_0(\mu) - \bar{\chi}_0(\nu)}{r-1} + \gamma\}$$



Together this implies

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{\bar{\chi}_0(\mu)}{r} + \gamma$$

(which is a refinement of Yomdin's global estimate by $\frac{R(f)}{r}$).

This implies

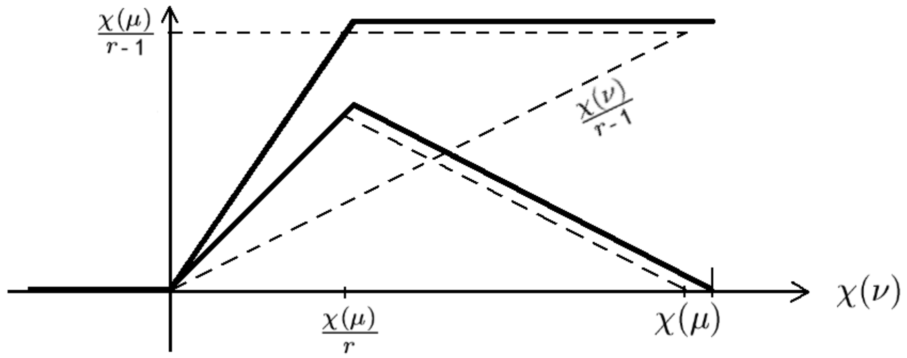
$$u_1(\mu) \le \frac{\bar{\chi}_0(\mu)}{r} \le \frac{\bar{\chi}_0(\mu)}{r-1}$$

Suppose

$$u_\alpha \le \frac{\bar{\chi}_0(\mu)}{r-1}$$

Then, near a measure μ

$$u_\alpha(\nu) + h^{New}(X|\nu,\mathcal{V}) \le \frac{\bar{\chi}_0(\nu)}{r-1} + \min\left\{\bar{\chi}_0(\nu), \frac{\bar{\chi}_0(\mu)-\bar{\chi}_0(\nu)}{r-1} + \gamma\right\} \le \frac{\bar{\chi}_0(\mu)}{r-1} + \gamma$$



which implies

$$u_{\alpha+1} \leq \frac{\bar{\chi}_0(\mu)}{r-1}$$

For limit ordinals the passage is trivial: if $u_\alpha \leq \frac{\bar{\chi}_0(\mu)}{r-1}$ for all $\alpha < \beta$ then

$$u_\beta = \widetilde{\sup_{\alpha<\beta} u_\alpha} \leq \frac{\bar{\chi}_0(\mu)}{r-1},$$

because the function on the right is u.s.c.

Eventually, $u_\alpha \leq \frac{\bar{\chi}_0(\mu)}{r-1}$ for all ordinals including α_0.

Using the transfinite formula we get the desired result:

$$h_{sex}(\mu) = h(\mu) + u_{\alpha_0}(\mu) \leq h(\mu) + \frac{\bar{\chi}_0(\mu)}{r-1}. \quad \square$$

- - - D A Y 2 - - -

PROOF OF THE ANTARCTIC THEOREM

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{\chi_0(\mu) - \chi_0(\nu)}{r-1} + \gamma,$$

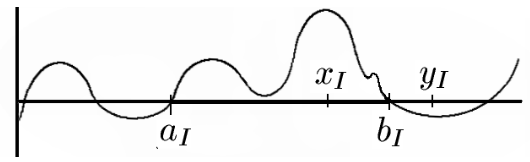
(4.1) Lemma. *Let $g : [0,1] \to \mathbb{R}$ be a C^r function, where $r > 0$. Then there exists a constant $c > 0$ such that for every $0 < s < 1$ the number of components of the set $\{x : g(x) \neq 0\}$ on which $|g|$ reaches or exceeds the value s is at most $c \cdot s^{-\frac{1}{r}}$.*

Proof. If g has a constant sign then there is only one component and the lemma holds with $c = 1$. Otherwise we proceed inductively, as follows: For $0 < r \leq 1$, g is Hölder, i.e., there exists a constant $c_1 > 0$ such that $|g(x) - g(y)| \leq c_1|x - y|^r$. If $|g(x)| \geq s$ and y is a zero point for g then

$$|x - y| \geq c_1{}^{-\frac{1}{r}} \cdot s^{\frac{1}{r}}.$$

The component containing x is at least that long and the number of such components is at most $c \cdot s^{-\frac{1}{r}}$, where $c = c_1{}^{\frac{1}{r}}$.

Now take $r > 1$ and suppose that the lemma holds for $r-1$. Let g be of class C^r. We count the components $I = (a_I, b_I)$ of $\{x : g(x) \neq 0\}$ where $|g|$ exceeds s. Unless $a_I = 0$ or $b_I = 1$, I contains a critical point. Let x_I denote the largest critical point $x \in I$ satisfying $|g(x)| \geq s$. Unless I is the last or last but one component, there is a critical point larger than or equal to b_I. Let y_I be the smallest such critical point. So, except for at most three components, I determines an interval (x_I, y_I).



Notice that these intervals are disjoint for different I. There are two possible cases: either
a) $y_I - x_I > s^{\frac{1}{r}}$, or
b) $y_I - x_I \leq s^{\frac{1}{r}}$.
Clearly, the number of components I satisfying a) is smaller than $s^{-\frac{1}{r}}$. If a component satisfies b) then, by the mean value theorem, $|g'|$ attains on (x_I, b_I) a value at least $s/s^{\frac{1}{r}} = s^{\frac{r-1}{r}}$. This value is attained on a component of the set $\{x : g'(x) \neq 0\}$ contained in (x_I, y_I). Because g' is of class C^{r-1}, by the inductive assumption, the number of such intervals (x_I, y_I) (hence of components I satisfying b)) does not exceed $c \cdot (s^{\frac{r-1}{r}})^{-\frac{1}{r-1}} = c \cdot s^{-\frac{1}{r}}$. Jointly, the number of all components I is at most $3 + (c+1) \cdot s^{-\frac{1}{r}} \leq (c+4) \cdot s^{-\frac{1}{r}}$. $\square$

Letting $g = f'$ we obtain the following

(4.2) Corollary. *Let $f : [0,1] \to [0,1]$ be a C^r function, where $r > 1$. Then there exists a constant $c > 0$ such that for every $s > 0$ the number of branches of monotonicity of f on which $|f'|$ reaches or exceeds s is at most $c \cdot s^{-\frac{1}{r-1}}$.*

(4.3) Definition. Let f be as in the formulation of Corollary (4.2). Let $\mathcal{I} = (I_1, I_2, \ldots, I_n)$ be a finite sequence of branches of monotonicity of f, (i.e., any formal finite sequence whose elements belong to the countable set of branches, admitting repetitions). Denote

$$a_i = \min\{-1, \max\{\log|f'(x)| : x \in I_i\}\}. \tag{4.4}$$

Choose $S \leq -1$. We say that $\mathcal{I}$ *admits* the value S if

$$\frac{1}{n}\sum_{i=1}^{n} a_i \geq S. \tag{4.5}$$

Clearly, if there exists a sequence of points $y_i \in I_i$ with $\log|f'(y_i)| \leq -1$ for each i and satisfying $\frac{1}{n}\sum_{i=1}^{n} \log|f'(y_i)| \geq S$, then $\mathcal{I}$ admits the value S.

For $t \in (0,1)$ we will use the notation $H(t) = -t\log t - (1-t)\log(1-t)$. Recall that this positive function approaches zero both at 0 and at 1. The standard application of Stirling's formula yields that for $m \geq n$ the logarithm of the binomial coefficient $\binom{m}{n}$ is bounded above by $mH(\frac{n}{m}) + 1$.

(4.6) Lemma. *Let $f : [0,1] \to [0,1]$ be a C^r function, where $r > 1$. Fix $\gamma > 0$. Then there exists $S_\gamma \leq -1$ such that for every n and $S < S_\gamma$ the logarithm of the number of sequences $\mathcal{I}$ of length n which admit the value S is at most*

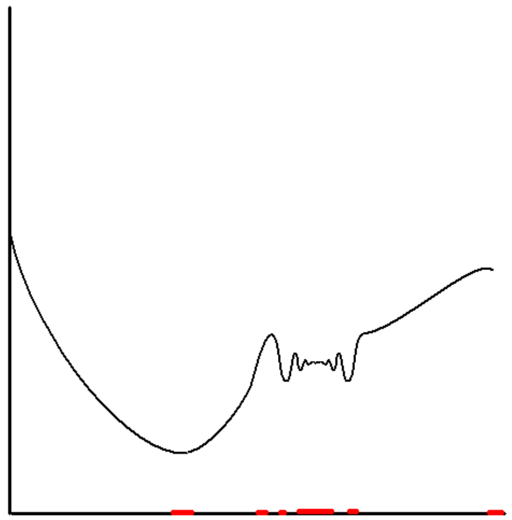
$$n\frac{-S}{r-1}(1+\gamma). \tag{4.7}$$

Proof. Without loss of generality assume that S is a negative integer. Let $\mathcal{I}$ be a sequence of n branches of monotonicity which admits the value S. Denote $k_i = \lfloor a_i \rfloor$. Then $(-k_i)$ is a sequence of n positive integers with sum at most $n(1-S)$. The number of such sequences (k_i) is bounded above by $\binom{n(1-S)}{n}$, and the logarithm of this number is dominated by $n(1-S)H(\frac{1}{1-S})+1$. Now, in a given sequence (k_i), each value k_i may be realized by any branch of monotonicity on which $\max\log|f'|$ lies between k_i and k_i+1 (or just exceeds -1 if $k_i = -1$). From Corollary (4.2) it follows that there are no more than $ce^{\frac{-k_i}{r-1}}$ such branches. Jointly the logarithm of the number of sequences of branches of monotonicity corresponding to one sequence (k_i) is at most $n\log c - \frac{1}{r-1}\sum_{i=1}^n k_i \leq n\log c + \frac{n}{r-1}(1-S)$, and the logarithm of the number of all sequences of branches of monotonicity which admit the value S is at most

$$n\log c + \tfrac{n}{r-1}(1-S) + n(1-S)H(\tfrac{1}{1-S}) + 1.$$

If S is close enough to minus infinity then the last expression does not exceed $n\frac{-S}{r-1}(1+\gamma)$ for any n. $\square$

Let us return to the transformation T of the interval or of the circle. In both cases the derivative f' of the associated function f can be regarded as a function defined on the interval $[0,1]$. Let $C = \{x : f'(x) = 0\}$ be the critical set. Fix $\gamma > 0$. Fix some open neighborhood U of C on which $\log|f'| < S_\gamma$. Notice that U^c can be covered by finitely many open intervals on which f is monotone. Let $\mathcal{V}$ be the cover consisting of U and these intervals.



(4.8) Lemma. *Let T be a C^r transformation of the interval or of the circle X, where $r > 1$. Let U and $\mathcal{V}$ be as described above. Let ν be an ergodic measure and let*

$$S(\nu) = \int_U \log |f'|\, d\nu. \tag{4.9}$$

Then

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{-S(\nu)}{r-1}(1+\gamma). \tag{4.10}$$

Proof. It suffices to consider the case of $S(\nu)$ finite. For $\sigma < 1$ there exists $n_\sigma \in \mathbb{N}$ such that the set F has measure larger than σ, where F is the set of points y satisfying, for every $n \geq n_\sigma$, that the nth Cesaro mean at y of the function $1_U \log |f'|$ equals $S(\nu)$ up to the error $1 - \sigma$. Let $x \in F$ and $n \geq n_\sigma$. Consider a set

$$V_x^n = V_0 \cap T^{-1}(V_1) \cap \cdots \cap T^{-n+1}(V_{n-1}) \tag{4.11}$$

containing x, with $V_i \in \mathcal{V}$ (as in the definition of local entropy). Consider the finite subsequence of times $0 \leq i_j \leq n-1$ when $V_{i_j} = U$. Let $n\zeta$ denote the length of this subsequence and assume $\zeta > 0$. For a fixed δ let E be an (n, δ)-separated set in $V_x^n \cap F$ and let $y \in E$. The sequence (i_j) contains only (usually not all) times i when $f^i(y) \in U$. Thus, since $y \in F$, we have

$$S(\nu) \leq \frac{1}{n}\Big(\sum_j \log|f'(T^{i_j}(y))| + A\Big) + 1 - \sigma, \tag{4.12}$$

where A is the similar sum over the times of visits to U not included in the sequence (i_j). Clearly $A \leq 0$, so it can be skipped. Dividing by ζ we obtain

$$\frac{S(\nu) - (1-\sigma)}{\zeta} \leq \frac{1}{n\zeta}\sum_j \log|f'(T^{i_j}(y))|. \tag{4.13}$$

The right hand side of (4.13) is smaller than S_γ. This implies that along the subsequence (i_j) the trajectory of y traverses a sequence $\mathcal{I}$ (of length $n\zeta$) of branches of monotonicity of f admitting the value $\frac{S(\nu)-(1-\sigma)}{\zeta}$ smaller than S_γ. By Lemma (4.6), the logarithm of the number of such sequences $\mathcal{I}$ is dominated by

$$n\frac{-S(\nu) + (1-\sigma)}{r-1}(1+\gamma). \tag{4.14}$$

At times i other than i_j the set V_i contains only one branch, so if two points from $V_x^n \cap F$ traverse the same sequence of branches along the times (i_j), they traverse the same full sequence of branches along all times $i = 0, 1, \ldots, n-1$ (this takes care also of the case when $\zeta = 0$). Now, it is easy to see that the number of (n, δ)-separated points which, along all times $i = 0, 1, \ldots, n-1$, traverse the same given sequence of branches of monotonicity is bounded above by $\frac{n}{\delta}$. This, together with (4.14), implies that the logarithm of the cardinality of E is at most

$$n \frac{-S(\nu) + (1-\sigma)}{r-1}(1+\gamma) + \log n - \log \delta. \tag{4.15}$$

The proof is concluded by dividing by n, letting $n \to \infty$ and then letting $\sigma \to 1$. $\square$

Proof of the Antarctic Theorem. Fix an invariant measure μ and some $\gamma > 0$. We need to consider only ergodic measures ν close to μ. If $\chi(\mu) < 0$ then, by upper semicontinuity of the function χ, for ν sufficiently close to μ, $\chi(\nu) < 0$, so by the Ruelle inequality (and since always $h^{New}(X|\nu, \mathcal{V}) \le h(\mu)$), $h^{New}(X|\nu, \mathcal{V}) = 0$ and the assertion (3.2) holds.
Now suppose that $\chi(\mu) \ge 0$. Clearly, then $\mu(C) = 0$. Since $\log|f'|$ is μ-integrable, for any given γ_1 the open neighborhood U of C (on which $\log|f'| < S_{\gamma_1}$) can be made so small that

$$\int_{\overline{U}} \log|f'(x)|\, d\mu > -\gamma_1. \tag{4.16}$$

Then

$$\int_{\overline{U}^c} \log|f'(x)|\, d\mu < \chi(\mu) + \gamma_1. \tag{4.17}$$

We define the cover $\mathcal{V}_\mu$ as $\mathcal{V}$ with the above choice of the set U (the parameter γ_1 will be specified at the end of the proof). The integral in (4.17) is an upper semicontinuous function of the measure ($\overline{U}^c$ is an open set on which $\log|f'|$ is finite and continuous and negative on the boundary). Thus there exists $\epsilon_\mu > 0$ such that $\mathrm{dist}(\nu, \mu) < \epsilon_\mu$ implies

$$\int_{\overline{U}^c} \log|f'(x)|\, d\nu < \chi(\mu) + \gamma_1. \tag{4.18}$$

The more

$$\int_{U^c} \log|f'(x)|\, d\nu < \chi(\mu) + \gamma_1 \tag{4.19}$$

(we have included the boundary to the set of integration, and the function is negative on that boundary). Then

$$-S(\nu) = \int_{U^c} \log |f'(x)|\, d\nu - \chi(\nu) \leq \chi(\mu) - \chi(\nu) + \gamma_1. \tag{4.20}$$

Substituting (4.20) into (4.10) we get

$$h^{New}(X|\nu, \mathcal{V}_\mu) \leq \frac{\chi(\mu) - \chi(\nu) + \gamma_1}{r-1}(1+\gamma_1). \tag{4.21}$$

Of course, $\chi(\mu)$ can be replaced by a not smaller number $\chi_0(\mu)$. If $\chi(\nu) < 0$ then $h^{New}(X|\nu, \mathcal{V}_\mu) = 0 \leq \frac{\chi_0(\mu)-\chi_0(\nu)}{r-1}$, so, in any case we can write

$$h^{New}(X|\nu, \mathcal{V}_\mu) \leq \frac{\chi_0(\mu) - \chi_0(\nu) + \gamma_1}{r-1}(1+\gamma_1). \tag{4.22}$$

Because $\frac{\chi_0(\mu)-\chi_0(\nu)}{r-1}$ is bounded above (for example by $\frac{\mathbf{L}(T)}{r-1}$), given $\gamma > 0$ we can choose γ_1 so small that the assertion (3.2) holds. $\square$

- - - D A Y 3 - - -

PROOF OF THE PASSAGE THEOREM

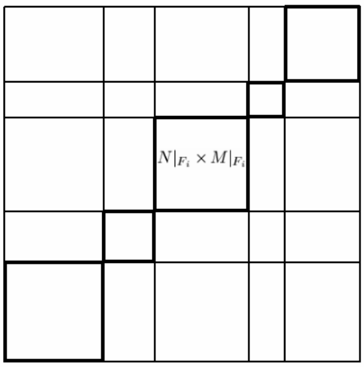
For every $\mu \in \mathcal{P}_T(X)$ and $\gamma > 0$ there exist an open cover $\mathcal{V}$ such that for ANY $\nu \in \mathcal{P}_T(X)$ sufficiently close to μ,

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{\bar{\chi}_0(\mu) - \bar{\chi}_0(\nu)}{r-1} + \gamma,$$

where $\bar{\chi}_0(\mu)$ is the average of χ_0 over the ergodic decomposition of μ.

(4.26) Lemma. *Let $\mathcal{P}$ be a compact metric space (with a metric* dist*) and let $\mathcal{M}$ be the set of all probability measures on $\mathcal{P}$ endowed with the weak* topology given by a metric* Dist. *Fix some $M \in \mathcal{M}$ and $\epsilon > 0$. Then there exists $\xi > 0$ such that for any $N \in \mathcal{M}$, $\mathrm{Dist}(N, M) < \xi$ implies that there exists a joining J of N and M (i.e., a measure on the product $\mathcal{P} \times \mathcal{P}$ with marginals N and M) such that $J(\Delta^\epsilon) > 1 - \epsilon$, where $\Delta^\epsilon = \{\langle \nu, \tau \rangle \in \mathcal{P} \times \mathcal{P} : \mathrm{dist}(\nu, \tau) < \epsilon\}$.*

Proof. There exists a partition of $\mathcal{P}$ into finitely many Borel sets F_i with the following properties: $\mathrm{diam}(F_i) < \epsilon$ and $M(\partial F_i) = 0$, for every i. (Here ∂F denotes the boundary of a set F.) Then every $N \in \mathcal{M}$ sufficiently close to M satisfies $\sum_i |N(F_i) - M(F_i)| < \epsilon$. For each i let $\alpha_i = \min\{N(F_i), M(F_i)\}$ and let J' be the subprobabilistic measure obtained as the sum of the product measures $N|_{F_i} \times M|_{F_i}$ normalized so that $J'(F_i \times F_i) = \alpha_i$. The marginals of J' are subprobabilistic measures M' and N' such that $M - M'$ and $N - N'$ are positive with equal masses β not exceeding ϵ. The joining J is obtained as the sum of J' and of $(M - M') \times (N - N')$ normalized to have the mass β. $\square$



(4.27) Corollary. *In a topological dynamical system* (X, T)*, let* $\mu, \nu_n \in \mathcal{P}_T(X)$*, and* $\nu_n \to \mu$ *in the weak* topology. Choosing a subsequence we can assume that* M_{ν_n} *converge to some* M*. By continuity of the barycenter map,* $\mathrm{bar}(M) = \mu$*. Then, given any* $\epsilon > 0$*, for* n *large enough, there exists a joining* J *of* M_{ν_n} *and* M *such that* $J(\Delta_e^\epsilon) > 1 - \epsilon$*, where*

$$\Delta_e^\epsilon = \{\langle \nu, \tau \rangle \in \mathcal{P}_T(X) \times \mathcal{P}_T(X) : \nu \textit{ is ergodic and } \mathrm{dist}(\nu, \tau) < \epsilon\}. \tag{4.28}$$

(The added condition that ν is ergodic is satisfied since each measure M_{ν_n} is, by definition, supported by the set of ergodic measures.)

Proof of the Passage Theorem. Suppose that there exists $\gamma > 0$ and a sequence ν_n converging to μ, and which, for any choice of an open cover $\mathcal{V}$ and $\epsilon > 0$, eventually does not satisfy the assertion (4.25) of the theorem. By choosing a subsequence we can assume that $M_{\nu_n} \to M$ with $\mathrm{bar}(M) = \mu$. Let γ_1 be such that $\gamma_1(1 + 3\mathbf{h}) \le \gamma$. For every τ in the support of M there is some open cover $\mathcal{V}_\tau$ and $\epsilon_\tau > 0$ established in the assumption applied to γ_1 and τ, (so that (4.24) is fulfilled with γ replaced by γ_1 and μ replaced by τ). For each τ the *Lebesgue number* of $\mathcal{V}_\tau$ is a positive number ξ_τ. Let $\epsilon < \gamma_1^2$ be so small that $\epsilon_\tau > \epsilon$ and $\xi_\tau > \epsilon$ for all τ belonging to a set G_1 of M-measure larger than $1 - \gamma_1$. We let $\mathcal{V}$ be an open cover by sets of diameter smaller than ϵ. This cover is finer than $\mathcal{V}_\tau$ for each $\tau \in G_1$, hence (4.24) holds for such τ, $\mathcal{V}$ and ϵ. By Corollary (4.27), for n large enough there exists a joining J of M_{ν_n} and M satisfying $J(\Delta_e^\epsilon) > 1 - \epsilon$. Let J_τ be the conditional probability measure of J with τ fixed on the second coordinate, and let ν_τ denote $\mathrm{bar}(J_\tau)$. We have

$$\int \nu_\tau \, dM(\tau) = \nu_n. \tag{4.29}$$

Recall that for almost every τ, J_τ is supported by ergodic measures ν. Moreover, for a set G_2 of M-measure at least $1 - \sqrt{\epsilon} > 1 - \gamma_1$ of τ's, J_τ is up to $\sqrt{\epsilon} < \gamma_1$ supported by the ϵ-neighborhood of τ. These conditions together imply that for $\tau \in G_1 \cap G_2$ (of M-measure at least $1 - 2\gamma_1$) all but a set of J_τ-measure γ_1 of measures ν are ergodic and so close to τ that they satisfy (4.24) with the parameters $(\gamma_1, \mathcal{V}, \epsilon, \nu, \tau)$ in place of $(\gamma, \mathcal{V}_\mu, \epsilon_\mu, \nu, \mu)$. Since local entropy is (by definition) harmonic, integrating both sides of (4.24) with respect to J_τ we obtain (for $\tau \in G_1 \cap G_2$)

$$h^{New}(X|\nu_\tau, \mathcal{V}) \le \frac{g_0(\tau) - \bar{g}_0(\nu_\tau)}{r - 1} + \gamma_1 + \gamma_1 \mathbf{h} \tag{4.30}$$

(on the "bad" set of ν's the local entropy is estimated above by $\mathbf{h}$). Now we integrate both sides of (4.30) over τ with respect to M. The term $g_0(\tau)$ will integrate to not more than the maximum over all measures M with barycenter μ. Such maximum was denoted as $\hat{g}_0(\mu)$. But recall (see Fact (2.5), in particular (2.8)) that this maximum is achieved at M_μ and equals $\bar{g}_0(\mu)$. The function $\bar{g}_0$ is harmonic, so, by (4.29), the term $\bar{g}_0(\nu_\tau)$ will integrate to $\bar{g}_0(\nu_n)$. Similarly will behave the left hand side. In this manner we obtain

$$h^{New}(X|\nu_n, \mathcal{V}) \leq \frac{\bar{g}_0(\mu) - \bar{g}_0(\nu_n)}{r-1} + \gamma_1 + \gamma_1\mathbf{h} + 2\gamma_1\mathbf{h} \tag{4.31}$$

(again, on the complement of $G_1 \cap G_2$ local entropy is estimated above by $\mathbf{h}$). By the choice of γ_1 we have contradicted the assumption that ν_n eventually does not satisfy the assertion (4.25) for γ, $\mathcal{V}$ and ϵ. $\square$