

Symbolic extensions of smooth interval maps*

Tomasz Downarowicz

*Institute of Mathematics and Computer Science,
Technical University, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: downar@im.pwr.wroc.pl*

Abstract: In this course we will present the full proof of the fact that every smooth dynamical system on the interval or circle X , constituted by the forward iterates of a function $f : X \rightarrow X$ which is of class C^r with $r > 1$, admits a symbolic extension, i.e., there exists a bilateral subshift (Y, S) with Y a closed shift-invariant subset of $\Lambda^{\mathbb{Z}}$, where Λ is a finite alphabet, and a continuous surjection $\pi : Y \rightarrow X$ which intertwines the action of f (on X) with that of the shift map S (on Y). Moreover, we give a precise estimate (from above) on the entropy of each invariant measure ν supported by Y in an optimized symbolic extension. This estimate depends on the entropy of the underlying measure μ on X , the “Lyapunov exponent” of μ (the genuine Lyapunov exponent for ergodic μ , otherwise its analog), and the smoothness parameter r . This estimate agrees with a conjecture formulated in [15] around 2003 for smooth dynamical systems on manifolds.

AMS 2000 subject classifications: Primary 37E05, 37C40; secondary 37A35.

Keywords and phrases: Entropy, interval maps, symbolic extensions.

Received April 2010.

Contents

1	Generally on symbolic extensions	84
2	The history of research on topological symbolic extensions	89
3	Introduction to entropy structures	92
4	The Newhouse entropy structure	93
5	Key ingredients in the one-dimensional result	93
6	Sketch of the proof of the Antarctic Theorem	96
7	Sketch of the proof of the Passage Theorem	100
	References	103

1. Generally on symbolic extensions

Nowadays, nearly every kind of information is turned into the digital form. Due to digital cameras and scanners, practically every image is turned into a computer file. The same happens to musical recordings or movies. Weather

*This article follows the lecture notes of an invited course held in December 2008 at DIM, University of Chile, Santiago. The subject is based on joint work with Alejandro Maass.

forecasting is done by complicated software which must be fed information in the digital form. Modern black boxes that register the history of airplane flights or truck rides do it in the digital form. Even our mathematical work is registered mainly as computer files. Analog information is nearly an extinct form.

While studying dynamical systems (in any understanding of this term) sooner or later one is forced to face the following question: “How can the information about the evolution of a given dynamical system be most precisely turned into a digital form?” As researchers specializing in dynamical systems, we are responsible for providing the theoretical background for such a transition.

So suppose that we are observing a dynamical system, and that we are indeed turning our observation into the digital form. That means, from time to time, we produce a digital “report”, a computer file, containing all our observations since the last report. Suppose for simplicity, that such reports are produced at equal time distances, say at integer times. Of course, due to bounded capacity of our recording devices and limited time between the reports, our files have bounded size (in bits). Because the variety of digital files of bounded size is finite, we can say, that at every integer moment of time we produce just one *symbol*, where the collection of all possible symbols (called the *alphabet* and denoted by Λ) is finite.

An illustrative example is filming a scene using a digital camera. Every unit of time, the camera registers an image, which is in fact a bitmap of some fixed size (camera resolution). The camera turns the live scene into a sequence of bitmaps. If the scene is filmed with sound, each bitmap is enhanced by a small sound file, also of bounded size. We can treat every such enhanced bitmap as a single symbol in the alphabet of the “language” of the camera.

The sequence of symbols is produced as long as the observation is being conducted. We have no reasons to restrict the global observation time, and we can agree that it goes on for ever. Sometimes (but not necessarily), we can also admit that the observation has been conducted since ever in the past as well. In this manner, the history of our recording takes on the form of a unilateral or bilateral sequence of symbols from some finite alphabet Λ . Advancing in time by a unit corresponds, on one hand, to the unit-time evolution of the dynamical system, on the other, to shifting the enumeration of our sequence of symbols. In this manner we have come to the conclusion, that the digital form of the observation is nothing else, but an element of the symbolic space $\Lambda^{\mathbb{S}}$, where \mathbb{S} stands either for the set of all integers \mathbb{Z} or nonnegative integers \mathbb{N}_0 . The action on this space is the familiar shift transformation σ given by $\sigma(x) = y$, where $x = (x_n)_{n \in \mathbb{S}}$ and $y = (x_{n+1})_{n \in \mathbb{S}}$.

Now, in most situations, such observation of the dynamical system will be *lossy*, i.e., it will capture only some aspects of the observed dynamical system. Much of the dynamics will be lost. For example, the digital camera will not be able to register objects hidden behind other objects, moreover, it will not see objects smaller than one pixel or their movements until they pass from one pixel to another. However, it may happen, that after a while, each object will eventually become visible, and that we will be able to reconstruct its trajectory from the recorded information.

Of course, lossy digitalization is always possible and hence presents a lesser kind of challenge. We will be much more interested in *lossless* digitalization. When is it possible to digitalize a dynamical system, so that no information is lost, i.e., in such a way, that after viewing the entire sequence of symbols, we can reconstruct the trajectory of every smallest particle in the system?

Well, it is certainly so, when the dynamical system under observation is not too complicated. When its rigidly moving particles are few, large, and the motion between the integer time moments is fully determined by the positions at the integer moments, and, at such moments each particle has only finitely many available positions. In other words, when the system is *discrete* in every aspect. But is this the only case?

The answer is no. At least at the purely theoretical level, the variety of systems that allow lossless digitalization is much larger. The class depends on the kind of approach we assume. We will concentrate on two levels: measure-theoretic and topological. Assuming the measure-theoretic point of view, each discrete time dynamical system is the action of a measure-preserving transformation on a measure space. We do not care about distances between particles, all we care about is partitions and probabilities with which the particles occupy the cells of these partitions. Here we are completely settled within the realm of ergodic theory. Assuming the topological point of view we do care about distances, but only up to preservation of convergence, i.e., we respect open and closed sets. In this setup we are within the realm of topological dynamics.

In the first, ergodic theoretic context, the question about “lossless digitalizability” of a system is relatively easy to answer. For automorphisms of probability spaces it is completely solved by the celebrated Krieger’s Generator Theorem: an automorphism T of a probability space (X, \mathcal{F}, μ) is isomorphic to the shift of a symbolic space $\Lambda^{\mathbb{Z}}$ (equipped with some shift-invariant measure) if and only if the Kolmogorov–Sinai entropy $h_{\mu}(T)$ of the automorphism is finite.

For endomorphisms, although the theorem no longer applies (in full generality) we can employ the notion of *natural extension*. If T is an endomorphism of a probability space, and has finite entropy, then its natural extension is an automorphism and has the same finite entropy. By the Krieger Theorem, this natural extension is isomorphic to a symbolic system. The original endomorphism becomes a *measure-theoretic factor* of the symbolic system. The natural extension in its digital (i.e., symbolic) form clearly contains complete information about all its factors, in particular, about the original endomorphism system, which, in this manner becomes losslessly digitalized.

On the other hand, any system (automorphism or endomorphism) of infinite entropy can be neither represented nor embedded in a symbolic system, because all symbolic systems have finite entropy. So, any digitalization of an infinite entropy system must be lossy. We have thus fully characterized measure-theoretic systems (on probability spaces) which are losslessly digitalizable: these are precisely the systems of finite Kolmogorov–Sinai entropy. The digitalization is then isomorphic either to the system itself, or, at worst, to its natural extension.

At the level of topological dynamics this problem is much more complicated. Here, given a topological dynamical system (X, T) (X is a compact metric

space and $T : X \rightarrow X$ is a continuous map, perhaps a homeomorphism), we seek for its digitalization in form of some, also topological, symbolic system. These are constituted by compact, shift-invariant subsets of the symbolic spaces $\Lambda^{\mathbb{S}}$ equipped with the action of the shift transformation σ . Such systems are shortly called *subshifts*. There are slight differences in our understanding of shift-invariance for the unilateral and bilateral sequences, but we skip these details here.

If we desire a symbolic system (subshift) (Y, σ) that carries all the information about a given topological dynamical system (X, T) , respecting its topological structure, a number of rather obvious limitations immediately pops out. First of all, we have very little chances to create a symbolic system that would be *topologically isomorphic* (i.e., *conjugate*) to (X, T) . Only *expansive* maps on *zero-dimensional* spaces are conjugate to subshifts. And these properties are rather exceptional among topological dynamical systems. In every other case we can only hope to build a *symbolic extension*, i.e., a subshift (Y, σ) , of which (X, T) would be a *topological factor*. There are equally little chances, that the extension will be conjugate to the topological natural extension of (X, T) . The natural extension would have to be zero-dimensional and expansive which implies that X is itself zero-dimensional and T nearly (not exactly but close to) expansive. So, the symbolic extension (Y, σ) , if one exist, will usually be something else than (X, T) or its natural extension. Such a (Y, σ) will contain other “unwanted” dynamics joined with the dynamics of (X, T) . It may even have necessarily larger topological entropy! Unlike in the measure-theoretic case, finite entropy (this time topological) does not even guarantee the existence of a symbolic extension. This is a phenomenon first discovered by Mike Boyle, whose interest in this subject was provoked by the question of Joe Auslander. Mike Boyle also indicated examples of systems with finite topological entropy, such that symbolic extensions do exist, but all have topological entropy essentially larger (by some constant) than that of (X, T) .

In this manner we are lead to studying the following general problem, which we can summarize in the two questions below, concerning a given topological dynamical system (X, T) .

QUESTION 1: *Does there exist a topological symbolic extension (Y, σ) of (X, T) ? In other words, is (X, T) a topological factor of some subshift?*

QUESTION 2: *If yes, what is the infimum of the topological entropies of all its symbolic extensions?*

These two questions (and some related ones) have triggered the creation of a relatively new branch in topological dynamics, the *theory of symbolic extensions*. It should not be surprising, that this theory is embedded in the theory of entropy of topological dynamical systems. In fact, it lead to some new developments in this theory, the discovery of some new entropy-related notions and invariants of topological conjugacy. It turns out, that in order to handle the two major questions posed above, one needs to focus not only on the topological entropies

of the involved systems (the system (X, T) and its symbolic extensions (Y, σ)), but also on the measure-theoretic (Kolmogorov–Sinai) entropies of all invariant measures supported by these systems. The two key notions of the theory are defined below.

Definition 1.1. Let (X, T) be a topological dynamical system. The *topological symbolic extension entropy* of (X, T) is defined as follows:

$$\begin{aligned} \mathbf{h}_{sex}(X, T) &= \inf\{\mathbf{h}_{top}(Y, \sigma) : (Y, \sigma) \text{ is a subshift extension of } (X, T)\}; \\ \mathbf{h}_{sex}(X, T) &= \infty \text{ if } (X, T) \text{ has no symbolic extensions.} \end{aligned}$$

A refinement of this notion at the level of invariant measures is provided below.

Definition 1.2. Let (X, T) be a topological dynamical system and let $\mathcal{P}_T(X)$ denote the set of all T -invariant measures μ on X . Let (Y, S) be a topological extension of (X, T) and $\pi : Y \rightarrow X$ be the corresponding factor map. On $\mathcal{P}_T(X)$ we define the *extension entropy function* by the formula

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

Then, on $\mathcal{P}_T(X)$ we define the *symbolic extension entropy function*, by

$$\begin{aligned} h_{sex}(\mu) &= \inf\{h^\pi(\mu) : \pi : Y \rightarrow X \text{ is a subshift extension of } (X, T)\}; \\ h_{sex}(\mu) &= \infty \text{ (for all measures } \mu) \text{ if } (X, T) \text{ has no symbolic extensions.} \end{aligned}$$

One of the fundamental tools in the theory of symbolic extensions is the following theorem (one inequality is obvious, the other requires some machinery):

Theorem 1.3 (Symbolic Extension Entropy Variational Principle).

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

The main task of the theory of symbolic extensions reduces to solving the following problem:

PROBLEM 1: *Compute (or estimate) h_{sex} for a given system (X, T) using its internal properties.*

Notice that the definition of h_{sex} is so constructed, that solving Problem 1 answers both of the formerly formulated questions 1 and 2. In full generality, so phrased problem has been solved in the paper [3], and then refined in [12]. The solution is in terms of so-called *entropy structure*, a carefully selected sequence of functions on $\mathcal{P}_T(X)$, which reflects the emergence of the entropy of different measures at refining scales. Crucial are upper semicontinuity properties of these functions and the multiple defect of uniformity in its convergence. The reason why these items are so essential can very roughly and briefly explained as follows: In the system (X, T) some invariant measures may reveal all of its entropy already in large scale (like in expansive systems), other measures may

need very small scale (i.e., fine covers) for their entropy to be detected. Now, in the symbolic extension (Y, σ) , the small scale dynamics must be “magnified” and become visible in the large scale of the symbolic system (in symbolic systems all dynamics happens in large scale). If the “large scale measures” are approximated in $\mathcal{P}_T(X)$ by the “small scale measures”, the magnification of small scale dynamics may lead to enlarging the entropy of large scale dynamics. This causes the overall entropy of the symbolic extension to grow.

In this course we will concentrate on a more particular problem, concerning smooth maps:

PROBLEM 2: *Estimate h_{sex} for C^r -maps on manifolds ($1 \leq r \leq \infty$).*

We will show how this problem is solved in dimension one, i.e., for smooth maps of the interval or of the circle, in terms of much more familiar parameters, such as the degree of smoothness r and the (slightly refined) Lipschitz constant.

2. The history of research on topological symbolic extensions

The first result concerning symbolic extensions in topological dynamics is due to William Reddy and goes back to 1968 ([21]). It says that every expansive homeomorphism T on a compact metric space has a symbolic extension. The construction provided no control over the entropy of this extension.

It was clear that expansiveness was a much too strong requirement. All known examples of finite entropy systems seemed to admit symbolic extensions. One of the spectacular applications of symbolic extensions occurs in the studies of hyperbolic systems. Using Markov partitions, such systems can be lifted to subshifts of finite type, which allows to apply symbolic dynamical methods to the hyperbolic systems. This approach belongs to the classics, it is described for example in Bowen’s book [1]. Generally, however, very little was known. The natural question whether all finite entropy systems indeed have symbolic extensions has been presumably puzzling many people between the years 1970 and 1990. Around 1989, Joe Auslander addressed this question to Mike Boyle, one of the best experts in symbolic dynamics. Within some time (less than a year), Boyle came up with the negative answer, by constructing an appropriate example. A version of the same example showed, that even if a system does admit a symbolic extension, there may exist a necessary gap between the entropy of the system and that of any symbolic extension. He called this gap *the residual entropy*. These examples have been presented at the Adler conference in 1991, but never published until 2002 (after the author of this note has already published his own version of Boyle’s examples in 2001). These examples proved only one thing: there is no easy answer to the questions 1 and 2 stated in the preceding section.

For the next 8 years, the progress was rather limited and not published. Mike Boyle collaborated in this matter with Doris and Ulf Fiebig. They tried to construct symbolic extensions by means of symbolic and topological methods

(without using invariant measures), which, from today’s perspective, explains why their results were so restricted.

Around 1998 the same problem was encountered by the author of this note. Together with Fabien Durand, they were characterizing all factors of so-called Toeplitz flows, and one of the three conditions for a system to be such a factor was that it admits some symbolic extension ([13]). It soon occurred, that nobody knew any general criteria for that. Mike Boyle was able to say that any system of entropy zero has a symbolic extension also of entropy zero ([2]), which was very useful for the study of factors of Toeplitz flows.

In year 1999, the author of this note spent a month in Marseille, devoting all his energy trying to understand why some systems have and other do not have symbolic extensions. For simplicity, he focused on zero dimensional systems, which seemed to be the best class to study. He discovered that the existence of symbolic extensions depends on the distribution of entropy on invariant measures, which lead to the first result containing the criteria for the existence and an estimate of the topological entropy of symbolic extensions for general zero-dimensional systems ([11]). In particular, he showed that an asymptotically h -expansive zero-dimensional system admits a symbolic extension of the same topological entropy. In the same paper he published the already mentioned examples based on those by Mike Boyle.

A year later, Boyle and the Fiebigers publish a long paper containing the results of their long lasting collaboration ([4]). The old examples appear here in the original version, next to new ones, where the transformation is on a disc and is differentiable at all but one point. In terms of positive results, all asymptotically h -expansive systems (not necessarily zero-dimensional) are shown to possess *principal* symbolic extensions, i.e., such that not only the topological entropy is the same as that of (X, T) , but also the Kolmogorov–Sinai entropy of every invariant measure is the same as that of its image in the system (X, T) . Since expansive systems are asymptotically h -expansive, we recover here a refined version of Reddy’s first result. Since any system of entropy zero is asymptotically h -expansive, we also recover the fact communicated earlier by Boyle to the author of this note. Another spectacular application, neatly included in [4] concerns smooth maps. Soon before that, Jerome Buzzi just proved that any C^∞ map on a Riemannian manifold is in fact asymptotically h -expansive ([9]). (Many years earlier Sheldon Newhouse proved a seemingly weaker statement [20], which from today’s perspective is equivalent to Buzzi’s result.) Now, this fact receives a new meaning: every C^∞ map on a manifold admits a principal symbolic extension. If we agree, that symbolic extensions are “lossless digitalizations”, then principal symbolic extensions can be regarded “gainless” (without superfluous information) digitalizations. The fact that all C^∞ maps can be losslessly and gainlessly digitalized became one of the iconic achievements of the theory of symbolic extensions. However, an immediate question arises: what about C^r maps, where $r < \infty$?

In 2001, the author of this note visited Mike Boyle. Leaving the smooth systems aside, they worked on the general theory. Their work [3] contains the complete and general characterization of the symbolic extension entropy func-

tion h_{sex} . It also contains the aforementioned variational principle for the symbolic extension entropy. Problem 1 and both questions stated in the preceding section, became completely solved. The solution still refers to zero-dimensional systems: each system with finite entropy is first shown to possess a principal zero-dimensional extension (using the theory of *mean dimension*, by E. Lindenstrauss and B. Weiss [17, 16]) and then it is shown how to build a symbolic extension of a zero-dimensional system. The notion of an *entropy structure* is introduced for zero-dimensional systems, the key tool to compute the symbolic extension entropy function. A criterion is provided, when the symbolic extension entropy function is *attained*, i.e., when a symbolic extension exist, whose entropy function matches the symbolic extension entropy function (an “optimal” digitalization).

Next year, the author of this work develops a consistent theory of entropy structures for general topological dynamical systems ([12]). Among other things, this allows to simplify the phrasing of several results from the preceding work, by skipping the intermediate stage of a zero-dimensional extension. The theory of entropy structures, although its importance depends upon the application to symbolic extensions, has gained an independent interest and several papers appeared devoted to other aspects of the entropy structure theory ([8, 18]).

At the same time the author collaborates with Sheldon Newhouse. The focus of this collaboration is on smooth maps on Riemannian manifolds. The obtained results ([15]) are of negative nature: roughly speaking they prove that (in some class) a typical C^1 system of dimension $d \geq 2$ admits no symbolic extensions at all (infinite symbolic extension entropy), while a typical C^r map, where $1 < r < \infty$ (also for $d \geq 2$) does not admit a principal symbolic extension (without saying that it does admit a symbolic extension). In their examples the gap between the entropy of the system and the entropy of a symbolic extension (the residual entropy) is bounded below by some term (which we denote here by \mathcal{R}) proportional to the Lipschitz constant and inverse proportional to $r - 1$. They formulate a conjecture, that the residual entropy in their examples is the worst, i.e., that every C^r map with $r > 1$ does admit a symbolic extension, and the symbolic extension entropy is, in the worst case, equal to the entropy plus \mathcal{R} .

This conjecture triggered a number of papers containing partial results. In all cases the conjecture has been confirmed: In 2005 the author of this note, jointly with Alejandro Maass, proves the conjecture true in dimension $d = 1$ ([14], the subject of this course). This result is then complemented by David Burguet, who provides examples of C^r interval maps showing the estimate \mathcal{R} for the residual entropy to be sharp ([5]). In the meantime Lorenzo Díaz and Todd Fisher prove related results for partially hyperbolic diffeomorphisms ([10]). Recently, Burguet proved the conjecture in two more cases: for C^r nonuniformly expanding maps (such that every invariant measure of positive entropy has all Lyapunov exponents nonnegative) on manifolds of any dimension and for any $r > 1$ ([6]), and, even more recently, for any C^2 surface diffeomorphisms ([7]). The general case of a C^r map (or diffeomorphism) on a compact manifold of dimension d remains an open problem, and the latest Burguet’s result for $d = 2$ is the most advanced step toward the full solution.

3. Introduction to entropy structures

For the purposes of this course, we will not need the general definition of the entropy structure. It suffices to know, that any entropy structure has the form of a sequence of functions $h_k : \mathcal{P}_T(X) \rightarrow [0, \infty)$, such that $h_k(\mu) \nearrow h(\mu)$ for every invariant measure μ . Sometimes it is better to consider the *tails* $\theta_k = h - h_k$. Then we have $\theta_k \searrow 0$ pointwise. Not all sequences $(\theta_k)_{k \geq 1}$ converging monotonically to zero are entropy structures. There are additional conditions on *how* they converge in reference to the dynamics. Still, there are many possible entropy structures in one dynamical system, but they are all equivalent to each other in a specific sense. Instead of listing the conditions which classify a given sequence (θ_k) as an entropy structure we will simply specify one particular such sequence (θ_k) , which has been proved to satisfy these conditions in the paper [12]. Only this entropy structure will be used throughout this course. The precise description of this sequence will be given in the next section.

So suppose we have already chosen an entropy structure (θ_k) . This allows to compute the symbolic extension entropy function h_{sex} . The derivation of h_{sex} from the entropy structure is via the “transfinite sequence”, as defined below:

$$\begin{aligned} \text{Step 0:} & \quad u_0 \equiv 0 \\ \text{Step } \alpha: & \quad v_\alpha = \sup_{\beta < \alpha} u_\beta, \quad u_\alpha = \lim_k \widetilde{v_\alpha + \theta_k}, \end{aligned}$$

(recall that $\tilde{f}(x) = \limsup_{y \rightarrow x} f(y)$).

Theorem 3.1 ([3]). *There exists a countable ordinal α_0 such that $u_\alpha = u_{\alpha_0}$ for every $\alpha \geq \alpha_0$, and*

$$h_{sex} = h + u_{\alpha_0}.$$

Combining this with Theorem 1.3 we get

$$\mathbf{h}_{sex}(T) = \sup_{\mu \in \mathcal{P}_T(X)} (h(\mu) + u_{\alpha_0}(\mu)).$$

As a digression, let us mention, that the theory of entropy structures allows to characterize the famous Misiurewicz parameter $\mathbf{h}^*(T)$ (the one used to define asymptotically h -expansive systems, by $\mathbf{h}^*(T) = 0$) as the pointwise supremum of the function u_1 :

$$\mathbf{h}^*(T) = \sup_{\mu \in \mathcal{P}_T(X)} u_1(\mu).$$

The two parameters appear at opposite poles of the transfinite sequence: $\mathbf{h}^*(T)$ is the “supremum of the first order”, while $\mathbf{h}_{sex}(T)$ is the “supremum of all orders”. The participation of $h(\mu)$ in only one of the suprema causes the two notions not to be related by any inequality. Only one implication holds in general: $\mathbf{h}^*(T) = 0 \implies \mathbf{h}_{sex}(T) = \mathbf{h}_{top}(T)$. In fact we have the equivalence:

$$\mathbf{h}^*(T) = 0 \text{ if and only if } \forall \mu \in \mathcal{P}_T(X) \ h_{sex}(\mu) = h(\mu).$$

This is to say, asymptotically h -expansive systems are exactly those which admit a principal symbolic extension ([4]).

4. The Newhouse entropy structure

Now we provide the definition of a local entropy, created by Sheldon Newhouse in 1989. Later, the author of this note has verified, that local entropy with respect to a refining sequence of open covers becomes an entropy structure. Below we use the following notation: F is any Borel subset of X , \mathcal{V} is an open cover of X and V_x^n is any set containing x and having the form $V_0 \cap T^{-1}(V_1) \cap \dots \cap T^{-n+1}(V_{n-1})$, where $V_0, V_1, \dots, V_{n-1} \in \mathcal{V}$. A set E is (n, δ) -separated if for any two points $y, y' \in E$ there is some $i \in \{0, 1, \dots, n-1\}$ with $d(T^i y, T^i y') \geq \delta$. Also μ denotes an invariant measure and σ is a number smaller than 1.

Definition 4.1 ([20]).

- (a) $H(n, \delta|x, F, \mathcal{V}) := \log \max\{\#E : E \text{ is an } (n, \delta)\text{-separated set in } F \cap V_x^n\}$;
- (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) $h(X|F, \mathcal{V}) := \lim_{\delta \rightarrow 0} h(\delta|F, \mathcal{V})$;
- (e) $h(X|\mu, \sigma, \mathcal{V}) := \inf\{h(X|F, \mathcal{V}) : \mu(F) > \sigma\}$;
- (f) for an ergodic measure μ , $h^{New}(X|\mu, \mathcal{V}) := \lim_{\sigma \rightarrow 1} h(X|\mu, \sigma, \mathcal{V})$.

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 ([12]).

5. Key ingredients in the one-dimensional result

In this section we state the main result of [14] and two key theorems leading to it. The first one, called “The Antarctic Theorem” is an estimate of local entropy for C^r interval (or circle) maps. The exotic name of the theorem comes from the fact that the breakthrough in proving it was made during the author’s trip to Antarctica, in fact while he was spending a sleepless night camping on the snow on one of the Antarctic islands. This is the only statement in this course, which uses the specific properties of the interval. The second intermediate result, called “The Passage Theorem” can be phrased as a completely general fact and in this form has already been used by Burguet in his two latest results. Its name reflects that it provides a “bridge” between the local entropy estimate of

the preceding theorem and the final estimate of the symbolic extension entropy function, given in the main result. One can also associate the name with Drake Passage, where, returning from Antarctica, the author attempted to apply his discovery to symbolic extensions (which was accomplished after returning to Santiago, with the help of the coauthor A. Maass). This section also contains the derivation of the Estimate Theorem from the two intermediate theorems.

The detailed proofs of the Antarctic, Passage and Estimate Theorems can be found in [14]. In this course, we sketch these proofs skipping some details. Instead, we will try to be more convincing by illustrating some of the arguments using figures.

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

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

(For ergodic measures in dimension one, $\chi(\mu)$ is the Lyapunov exponent.)

Theorem 5.1 (The Antarctic Theorem). *Fix some $\gamma > 0$. For each $\mu \in \mathcal{P}_f(X)$ there exists an open cover \mathcal{V} of X such that*

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

for every ergodic measure ν in an open neighborhood of μ in $\mathcal{P}_f(X)$.

The Passage Theorem says the same, but without assuming ergodicity of ν . The function $\bar{\chi}_0(\mu)$ is defined by averaging χ_0 over the ergodic decomposition of μ . Since χ_0 is evidently convex, $\bar{\chi}_0$ is usually slightly larger than χ_0 (except at ergodic measures, where these two are equal).

Theorem 5.2 (The Passage Theorem). *Fix some $\gamma > 0$. For each $\mu \in \mathcal{P}_f(X)$ there exists an open cover \mathcal{V} of X such that*

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

for every invariant measure ν in an open neighborhood of μ in $\mathcal{P}_f(X)$.

The main result is this:

Theorem 5.3 (The Estimate Theorem). *Let f be a C^r transformation of the interval or of the circle X , where $r > 1$. Then*

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

As a consequence, by the symbolic extension entropy variational principle (and the usual variational principle),

$$\mathbf{h}_{sex}(f) \leq \mathbf{h}_{top}(f) + \frac{L(f)}{r-1},$$

where $L(f)$ is the logarithm (if positive) of the Lipschitz constant of f , that is $L(f) = \log^+ \sup_{x \in X} |f'(x)|$.

Remark 5.4. The Lipschitz constant can be easily replaced by the smaller constant $R(f) = \lim_{n \rightarrow \infty} \frac{1}{n} L(f^n)$, where f^n denotes the composition power of f .

We now describe how the Estimate Theorem is deduced from the Passage Theorem. This is fairly easy.

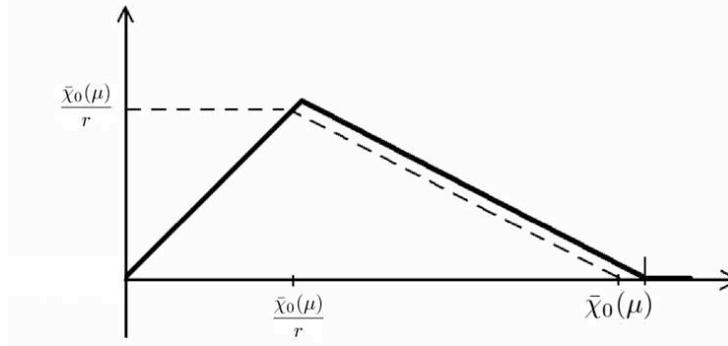
So, assume the Passage Theorem holds. The Ruelle inequality ($h(\nu) \leq \chi_0(\nu)$ for ergodic ν , see [22]) easily implies that for any invariant measure ν we also have

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

Thus, for ν sufficiently close to μ we have

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

Clearly, $h^{New}(X|\nu, \mathcal{V})$ (as well as $\bar{\chi}_0(\nu)$) cannot be negative. The situation is illustrated on the figure below. The horizontal axis represents all measures ν in the vicinity of μ parametrized by $\bar{\chi}_0(\nu)$, on the vertical axis we have the upper bound for $h^{New}(X|\nu, \mathcal{V})$:



It is seen from this picture (which replaces elementary calculations), that, for all considered measures ν ,

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

(which is a refinement of Yomdin's global estimate by $\frac{R(f)}{r}$, see [Y]). Plugging this to the definition of the transfinite sequence, we obtain

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

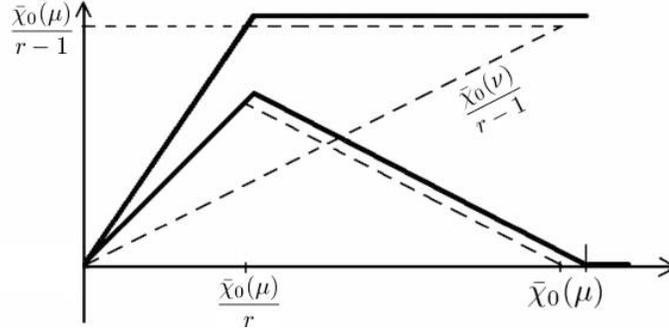
We proceed by the transfinite induction. Suppose

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

for all ordinals $\beta < \alpha$. Then, near a measure μ , holds

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

The situation is shown on the figure below.



We are using the fact that $\bar{\chi}_0$ is an upper semicontinuous function, hence in a sufficiently small vicinity of μ all measures ν satisfy $\bar{\chi}_0(\nu) \leq \bar{\chi}_0(\mu) + \gamma$. This is why the domain of the graph extends only a bit beyond $\bar{\chi}_0(\mu)$ (further to the right it would grow, so we are happy not have to include that part).

By passing to the upper limit as ν approaches μ , we get

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

By transfinite induction, $u_\alpha \leq \frac{\bar{\chi}_0(\mu)}{r-1}$ for all ordinals including α_0 . Now, using the transfinite characterization of the symbolic extension entropy, we get the desired result:

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

□

6. Sketch of the proof of the Antarctic Theorem

The proof relies on the following, fairly elementary counting lemma:

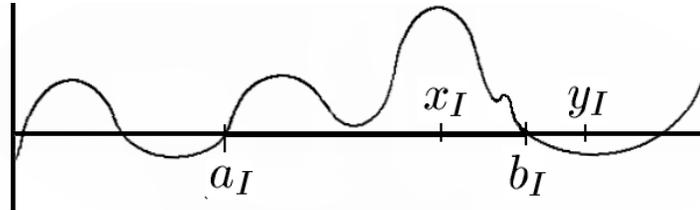
Lemma 6.1. *Let $g : [0, 1] \rightarrow \mathbb{R}$ be a C^r function, where $r > 0$. Then there exists a constant $c > 0$ such that for every $0 < s < 1$ the number of components of the set $\{x : g(x) \neq 0\}$ on which $|g|$ reaches or exceeds the value s is at most $c \cdot s^{-\frac{1}{r}}$.*

Proof. 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 $c = c_1^{\frac{1}{r}}$.

For larger r we proceed inductively: suppose that the lemma holds for $r - 1$. Let g be of class C^r . By elementary considerations of the graph of f , with each component $I = (a_I, b_I)$ of the set $\{x : g(x) \neq 0\}$ we can disjointly associate an interval (x_I, y_I) , so that $|g|$ attains at x_I its maximum on I and y_I is a critical point lying to the right of I (see the figure below).



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) and $|f|$ exceeds s on it, then, by the mean value theorem, $|g'|$ attains on (x_I, b_I) a value at least $s/s^{\frac{1}{r}} = s^{\frac{r-1}{r}}$. 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) 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 on which $|f|$ exceeds s is at most $2 + (c+1) \cdot s^{-\frac{1}{r}} \leq c_1 \cdot s^{-\frac{1}{r}}$ (the number 2 is added because the above argument does not apply to the extreme components, which need not contain critical points). \square

For $g = f'$ we obtain the following.

Corollary 6.2. *Let $f : [0, 1] \rightarrow [0, 1]$ 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 $c \cdot s^{-\frac{1}{r-1}}$.*

Next we apply the above to counting possible ways by which a point with a bounded below derivative for the composition power of f may traverse the branches of monotonicity. We make a formal definition.

Definition 6.3. Let f be as in the formulation of Corollary 6.2. Let $\mathcal{I} = (I_1, I_2, \dots, I_n)$ 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

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

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

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

Notice that, 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 .

Lemma 6.4. *Let $f : [0, 1] \rightarrow [0, 1]$ be a C^r function, where $r > 1$. Fix $\epsilon > 0$. Then there exists $S_\epsilon \leq -1$ such that for every n and $S < S_\epsilon$ the logarithm of the number of sequences \mathcal{I} of length n which admit the value S is at most*

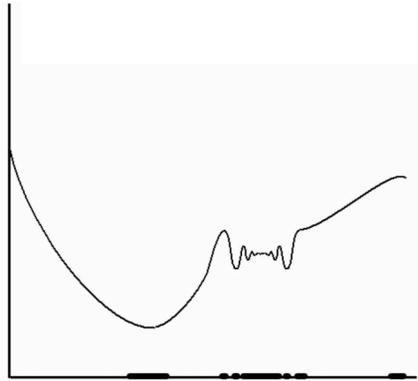
$$n \frac{-S}{r-1} (1 + \epsilon).$$

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 $n(1 - S)$. 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 6.2 it follows that there are no more than $ce^{\frac{-k_i}{r-1}}$ such branches for each k_i . Jointly the logarithm of the number of sequences of branches of monotonicity corresponding to one sequence (k_i) is at most

$$n \log c - \frac{1}{r-1} \sum_{i=1}^n k_i \leq n \log c + \frac{n}{r-1} (1 - S),$$

For large values of $-S$, the first term, and the last 1 can be skipped at a cost of multiplying $-S$ by $(1 + \epsilon)$. The number of all possible sequences (k_i) with sum $n(1 - S)$ is negligibly small on the logarithmic scale. So the logarithm of the number of all sequences of branches of monotonicity which admit the value S is, regardless of n , estimated from above as in the assertion. \square

Regardless of whether f is a transformation of the interval or of the circle X , the derivative f' can be regarded as a function defined on the interval $[0, 1]$. Let $C = \{x : f'(x) = 0\}$ be the critical set. Fix $\epsilon > 0$. Fix some open neighborhood U of C on which $\log |f'| < S_\epsilon$. Then 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. The figure below shows f and the set U .



Lemma 6.5. *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*

$$S(\nu) = \int_U \log |f'| d\nu.$$

Then

$$h^{New}(X|\nu, \mathcal{V}) \leq \frac{-S(\nu)}{r-1}(1+\epsilon). \quad (6.1)$$

Proof. Let F be the set of points on which the n th Cesaro means of the function $\mathbf{1}_U \log |f'|$ are close to $S(\nu)$ for n larger than some threshold integer (we are using the ergodic theorem; such a set F can have measure larger than $1-\sigma$). For $x \in F$ and large n consider a set

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

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

$$S(\nu) \approx \frac{1}{n} \left(\sum_j \log |f'(T^{i_j}(y))| + A \right),$$

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

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

The right hand side above is smaller than S_ϵ . 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 $\frac{S(\nu)}{\zeta}$ smaller than S_ϵ . By Lemma 6.4, the logarithm of the number of such sequences \mathcal{I} is dominated by

$$n \frac{-S(\nu)}{r-1} (1+\epsilon). \quad (6.2)$$

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 $i = 0, 1, \dots, n-1$. The number of (n, δ) -separated points which, along all times $i = 0, 1, \dots, n-1$, traverse the same given sequence of branches of monotonicity is negligibly small. This, together with (6.2), implies that the logarithm of the cardinality of E can be only negligibly larger than (6.2). The proof is concluded by dividing by n , and letting $n \rightarrow \infty$. \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 $h^{New}(X|\nu, \mathcal{V}) \leq h(\mu)$), $h^{New}(X|\nu, \mathcal{V}) = 0$ and the assertion holds for any open cover.

Now suppose that $\chi(\mu) \geq 0$. Clearly, then $\mu(C) = 0$. Since $\log |f'|$ is μ -integrable, the open neighborhood U of C (on which $\log |f'| < S_\epsilon$) can be made so small that the (negative) integral of $\log |f'|$ over the closure of U is very close to zero (closer than some ϵ). Then

$$\int_{\overline{U^c}} \log |f'(x)| d\mu < \chi(\mu) + \epsilon. \quad (6.3)$$

The integral in (6.3) 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), hence (6.3) holds for all invariant measures ν in a neighborhood of μ . The more

$$\int_{\overline{U^c}} \log |f'(x)| d\nu < \chi(\mu) + \epsilon$$

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

$$-S(\nu) = \int_{\overline{U^c}} \log |f'(x)| d\nu - \chi(\nu) \leq \chi(\mu) - \chi(\nu) + \epsilon. \quad (6.4)$$

We define the cover \mathcal{V} with the above choice of the set U (recall, \mathcal{V} consists of U and some intervals on which f is monotone). We can now apply Lemma 6.5. Substituting (6.4) into (6.1) we get

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

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

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

Because the function $\frac{\chi_0(\cdot)}{r-1}$ is bounded, the contribution of the error terms ϵ can be made smaller than the additive term γ . \square

7. Sketch of the proof of the Passage Theorem

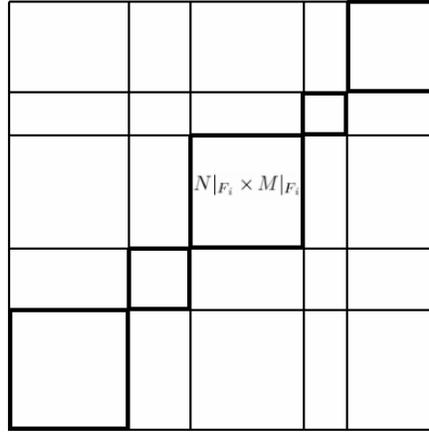
In the Passage Theorem, we need to drop the assumption that ν is ergodic. The key tool is the lemma below. Recall that the ergodic decomposition allows to represent each invariant measure ν as the barycenter of a probability measure M_ν supported by the set of ergodic measures. In order to easier distinguish

between probability measures on $\mathcal{P}_T(X)$ and invariant measures on X (which are points in $\mathcal{P}_T(X)$) we will consistently use the term “distribution” with regard to probability measures on $\mathcal{P}_T(X)$, in particular to the ergodic (and nonergodic) decompositions of invariant measures. Below, by a *joining* of two distributions M, M' on some space we understand any distribution on the Cartesian square of the space, with marginals M, M' .

Lemma 7.1. *In a topological dynamical system (X, T) , let $\mu, \nu_n \in \mathcal{P}_T(X)$, and $\nu_n \rightarrow \mu$ in the weak* topology. Choosing a subsequence we can assume that the ergodic decompositions M_{ν_n} converge to some distribution M on $\mathcal{P}_T(X)$. By continuity of the barycenter map, $\text{bar}(M) = \mu$. Then, given any $\epsilon > 0$, for n large enough, there exists a joining J_n of M_{ν_n} and M such that $J_n(\Delta_\epsilon^\epsilon) > 1 - \epsilon$, where*

$$\Delta_\epsilon^\epsilon = \{(\nu, \tau) \in \mathcal{P}_T(X) \times \mathcal{P}_T(X) : \nu \text{ is ergodic and } \text{dist}(\nu, \tau) < \epsilon\}.$$

Proof. The proof is elementary, and we only sketch it. We partition $\mathcal{P}_T(X)$ into finitely many Borel sets F_i of diameter smaller than ϵ and with boundaries of measure M zero. Then, for large n the numbers $M_{\nu_n}(F_i)$ are very close to $M(F_i)$ (for every index i). The joining J_n is obtained as the sum of appropriately scaled product distributions $M|_{F_i} \times M_{\nu_n}|_{F_i}$. Such a joining is supported by the ϵ -neighborhood of the diagonal (see figure below). □



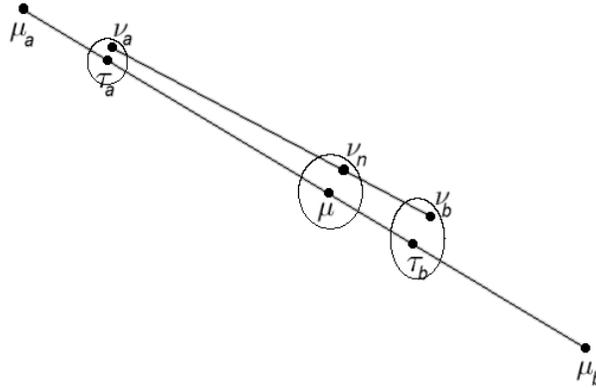
Proof of the Passage Theorem. Suppose that there exists $\gamma > 0$ and a sequence ν_n converging to μ , and which, for any choice of an open cover \mathcal{V} , eventually does not satisfy the assertion of the Passage Theorem. By choosing a subsequence we can assume that the ergodic decompositions M_{ν_n} converge to some distribution M on $\mathcal{P}_T(X)$ with $\text{bar}(M) = \mu$. By the Antarctic Theorem, for every τ in the support of M there is some open cover \mathcal{V}_τ and $\epsilon_\tau > 0$ such that

$$h^{New}(X|\tau, \mathcal{V}_\tau) \leq \frac{\chi_0(\mu) - \chi_0(\nu)}{r - 1} + \gamma, \tag{7.1}$$

for any ergodic ν in the ϵ_τ -neighborhood of τ . For each τ the *Lebesgue number* of \mathcal{V}_τ is a positive number ξ_τ . Let ϵ be so small that $\epsilon_\tau > \epsilon$ and $\xi_\tau > \epsilon$ for M -nearly all τ (belonging to a set $\mathcal{P} \subset \mathcal{P}_T(X)$ with $M(\mathcal{P}) \approx 1$). We let \mathcal{V} be an open cover by sets of diameter smaller than ϵ . This cover is finer than \mathcal{V}_τ for M -nearly each τ , hence (7.1) holds for such τ , \mathcal{V} and ϵ . By Lemma 7.1, for n large enough there exists a joining J_n of M_{ν_n} and M satisfying $J_n(\Delta_\epsilon^\epsilon) > 1 - \epsilon$. We fix such an n and let J_τ be the conditional of J_n for τ fixed on the second coordinate, and we let ν_τ denote $\text{bar}(J_\tau)$. We have

$$\int \nu_\tau dM(\tau) = \nu_n. \tag{7.2}$$

By the properties of the joining J_n , for M -nearly all τ the distribution J_τ is nearly supported by the ϵ -neighborhood of τ . These conditions together imply that for M -nearly every τ the distribution J_τ is nearly supported by the ergodic measures ν which satisfy (7.1) for the cover \mathcal{V} .



The idea of the above argument is presented on this figure. For simplicity, ν_n is shown as a convex combination of two ergodic measures ν_a and ν_b (the distribution M_{ν_n} is supported by these two points). The limit distribution M has barycenter μ and in this figure is also supported by two points τ_a and τ_b (not necessarily ergodic; M need not be the ergodic decomposition of μ , which in this case is a convex combination of completely different measures μ_a and μ_b). The role of the joining J_n is to associate to each τ in the support of M a “part” of the distribution M_{ν_n} , called J_τ , nearly supported by a small neighborhood of τ . In the case shown on the figure, it associates to τ_a the point mass at ν_a , and to τ_b the point mass at ν_b .

Integrating both sides of (7.1) with respect to J_τ we get for M -nearly every τ :

$$h^{New}(X|\nu_\tau, \mathcal{V}) \leq \frac{\chi_0(\tau) - \bar{\chi}_0(\nu_\tau)}{r - 1} + \gamma + \delta$$

(δ is the contribution of the “bad” measures ν). Now we integrate both sides of the above inequality over τ with respect to M . By convexity, $\chi_0(\tau)$ will integrate to not more than $\int \chi_0 dM_\mu$ which equals $\bar{\chi}_0(\mu)$. By (7.2), the term $\bar{\chi}_0(\nu_\tau)$ will integrate to $\bar{\chi}_0(\nu_n)$, while the left hand side will integrate to $h^{New}(X|\nu, \mathcal{V})$. Another negligibly small term representing the integral over the “bad” set of τ ’s will add. We have contradicted the assumption that ν_n eventually does not satisfy the assertion of the theorem. \square

References

- [1] BOWEN, R. (2008). *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, 2nd revised. Lect. Notes Math. **470**. Springer, Berlin. [MR2423393](#)
- [2] BOYLE, M. (1991) *Quotients of subshifts*. Adler Conference lecture.
- [3] BOYLE, M. and DOWNAROWICZ, T. (2004). The entropy theory of symbolic extensions. *Invent. Math.* **156**, no. 1, 119–161. [MR2047659](#)
- [4] BOYLE, M., FIEBIG, D. and FIEBIG, U. (2002). Residual entropy, conditional entropy and subshift covers. *Forum Math.* **14**, no. 5, 713–757. [MR1924775](#)
- [5] BURGNET, D. (2010). Examples of C^r interval map with large symbolic extension entropy. *Discrete and Continuous Dynamical Systems - A* **26**, no 3, 873–899.
- [6] BURGNET, D. (2009). Symbolic extensions for C^r nonuniformly entropy expanding maps. *Preprint*.
- [7] BURGNET, D. (2009). C^2 surface diffeomorphisms have symbolic extensions. *Preprint*.
- [8] BURGNET, D. and MCGOFF, K. (2010). Orders of accumulation of entropy. *Preprint*.
- [9] BUZZI, J. (1997). Intrinsic ergodicity of smooth interval maps. *Israel J. Math.* **100** 125–161. [MR1469107](#)
- [10] DÍAZ, L. and FISHER, T. (2009). Symbolic extensions for partially hyperbolic diffeomorphisms. *Preprint*.
- [11] DOWNAROWICZ, T. (2001). Entropy of a symbolic extension of a dynamical system. *Ergodic Theory Dynam. Systems* **21**, no. 4, 1051–1070. [MR1849601](#)
- [12] DOWNAROWICZ, T. (2005). Entropy structure. *J. Anal. Math.* **96** 57–116. [MR2177182](#)
- [13] DOWNAROWICZ, T. and DURAND, F. (2002). Factors of Toeplitz flows and other almost 1 – 1 extensions over group rotations. *Math. Scand.* **90**, no. 1, 57–72. [MR1887094](#)
- [14] DOWNAROWICZ, T. and MAASS, A. (2009). Smooth interval maps have symbolic extensions: the Antarctic theorem. *Invent. Math.* **176**, no. 3, 617–636. [MR2501298](#)
- [15] DOWNAROWICZ, T. and NEWHOUSE, S. (2005). Symbolic extensions and smooth dynamical systems. *Invent. Math.* **160**, no. 3, 453–499. [MR2178700](#)
- [16] LINDENSTRAUSS, E. (1999). Mean dimension, small entropy factors and an imbedding theorem. *Publ. Math. I.H.E.S.* **89** 227–262. [MR1793417](#)

- [17] LINDENSTRAUSS, E. and WEISS, B. (2000). Mean topological dimension. *Israel J. Math.* **115** 1–24. [MR1749670](#)
- [18] MCGOFF, K. (2010). Orders of accumulation of entropy on manifolds. *Preprint*.
- [19] MISIUREWICZ, M. (1976). Topological conditional entropy. *Studia Math.* **55**, no. 2, 175–200. [MR0415587](#)
- [20] NEWHOUSE, S. (1990). Continuity properties of entropy. *Ann. Math.* **129** 215–235. *Corr. in* **131** 409–410 (1990). [MR0986792](#)
- [21] REDDY, W. L. (1968). Lifting expansive homeomorphisms to symbolic flows. *Math. Systems Theory* **2** 91–92. [MR0224080](#)
- [22] RUELLE, D. (1978). An inequality for the entropy of differentiable maps. *Bol. Soc. Brasil. Mat.* **9**, no. 1, 83–87. [MR0516310](#)
- [23] YOMDIN, Y. (1987). Volume growth and entropy. *Israel J. Math.* **57**, no. 3, 285–300. [MR0889979](#)