THE ENTROPY THEORY OF SYMBOLIC EXTENSIONS

MIKE BOYLE AND TOMASZ DOWNAROWICZ

ABSTRACT. Fix a topological system (X, T), with its space $K(X, T)$ of T-invariant Borel probabilities. If (Y, S) is a symbolic system (subshift) and $\varphi : (Y, S) \to (X, T)$ is a topological extension (factor map), then the function h^{φ}_{ext} on $K(X, T)$ which assigns to each μ the maximal entropy of a measure ν on Y mapping to μ is called the *extension entropy function* of φ. The infimum of such functions over all symbolic extensions is called the *symbolic extension entropy function* and is denoted by h_{sex}. In this paper we completely characterize these functions in terms of functional analytic properties of an *entropy structure* on (X, T). The entropy structure $\mathcal{H}$ is a sequence of entropy functions h_k defined with respect to a refining sequence of partitions of X (or of $X \times Z$, for some auxiliary system (Z, R) with simple dynamics) whose boundaries have measure zero for all the invariant Borel probabilities. We develop the functional analysis and computational techniques to produce many dynamical examples; for instance, we resolve in the negative the question of whether the infimum of the topological entropies of symbolic extensions of (X, T) must always be attained, and we show that the maximum value of h_{sex} need not be achieved at an ergodic measure. We exhibit several characterizations of the *asymptotically h-expansive* systems of Misiurewicz, which emerge as a fundamental natural class in the context of the entropy structure. The results of this paper are required for the Downarowicz-Newhouse results [DN] on smooth dynamical systems.

CONTENTS

Date: July 31, 2003.
2000 *Mathematics Subject Classification.* Primary: 37B10; Secondary: 37B40, 37C40, 37C45, 37C99, 37D35.
Key words and phrases. sex entropy, residual entropy, conditional entropy, topological entropy, entropy, superenvelope, symbolic dynamics.
The research of the second named author was supported by a KBN grant.

1. Introduction

Suppose S, T are selfhomeomorphisms of compact metrizable spaces X, Y and $\varphi\colon Y \to X$ is a continuous surjection such that $T\varphi = \varphi S$. Then we refer to φ, and also (Y, S), as an extension of (X, T). We say it is a *symbolic extension* if (Y, S) is a subshift over a finite alphabet. Symbolic extensions are fundamental to the use of symbolic dynamics in the study of dynamical systems (e.g. [Bow]).

We introduce $h_{\text{sex}}(T)$, the *symbolic extension entropy* of T, by setting

$$\mathbf{h}_{\text{sex}}(T) = \inf\{\mathbf{h}_{\text{top}}(S)\colon (Y,S) \text{ is a symbolic extension of } (X,T)\}$$

where $\mathbf{h}_{\text{top}}(S)$ denotes the topological entropy of S. We adopt the convention that the infimum of the empty set is by definition ∞. Abbreviating, we use "sex entropy" as synonymous with "symbolic extension entropy". Sex entropy was studied [BFF, D2] in the notation of the residual entropy of T, which is by definition $\mathbf{h}_{\text{res}}(T) = \mathbf{h}_{\text{sex}}(T) - \mathbf{h}_{\text{top}}(T)$. The residual entropy gives a coarse measure of describability of a system by symbolic dynamics. The notational choice of sex entropy will be more natural for most of the current paper (e.g., Remark 8.2).

We define the extension entropy function of an extension $\varphi\colon (Y,S) \to (X,T)$ as

$$h^{\varphi}_{\text{ext}}\colon K(X,T) \to [0,\infty)$$
$$\mu \mapsto \sup\{h(S,\nu)\colon \nu \in K(Y,S) \text{ and } \varphi\nu = \mu\}$$

where e.g. $K(X,T)$ denotes the compact space of T-invariant Borel probabilities, $h(S,\nu)$ is the measure theoretic entropy of S with respect to the measure ν, and we let φ abbreviate the induced map φ_* on measures. For a given μ, the number $h^{\varphi}_{\text{ext}}(\mu)$ measures the complexity of the orbits appearing in Y above the support of μ. There is a variational principle ([LeWa, DS1]) which gives a precise a.e. formulation of this idea in terms of fiber entropies (see 6.5).

In this paper, we will study entropy jumps to symbolic systems at the level of measures. We introduce $h_{\text{sex}}(T,\cdot)$, the symbolic extension entropy function of T, by setting

$$h_{\text{sex}}(T,\mu) = \inf\{h^{\varphi}_{\text{ext}}(\mu)\colon \varphi \text{ is a symbolic extension of } T\}$$

and similarly we define the residual entropy function of T on $K(X,T)$ by $h_{\text{res}}(T,\mu) = h_{\text{sex}}(T,\mu) - h(T,\mu)$. Sex entropy as a function on measures provides a much finer probe into the dynamics of T than the number $\mathbf{h}_{\text{sex}}(T)$. We will obtain a rather complete functional analytic understanding of sex entropy functions.

For this, given a system (X,T), we define an entropy structure, generalizing to arbitrary systems the entropy structure introduced in [D2] in dimension zero as a key tool for studying residual entropy. First, suppose the system (X,T) admits a refining sequence of finite Borel partitions P_k with small boundaries (i.e., boundaries which have measure zero for every measure in $K(X,T)$). In this case (a common one – see 7.6), we may define an entropy structure for (X,T) to be the sequence $\mathcal{H}$ of functions $h_k\colon K(X,T) \to \mathbb{R}$ defined by the rule $h_k\colon \mu \mapsto h(T,\mu,P_k)$. For a general system (X,T), we define the sequence $\mathcal{H}$ with respect to $(X \times Z, T \times R)$, where (Z,R) can be any nonperiodic zero entropy minimal system (see 5.2). The point of this is that $(X \times Z, T \times R)$ admits the required refining sequence P_k with small boundaries, as a consequence of the deep constructions of Lindenstrauss [Li], using the mean dimension theory of dynamical systems developed by Lindenstrauss and Weiss [LiWe]. Our results will not depend on the particular sequence $\mathcal{H}$ chosen for a system (there is no canonical sequence). In any case $\mathcal{H}$ will be a nondecreasing sequence of affine nonnegative upper semicontinuous (u.s.c.) functions with u.s.c. differences.

We define a *superenvelope* of the sequence $\mathcal{H}$ to be a function E such that for each k the difference $E - h_k$ is nonnegative and u.s.c. We allow just one unbounded superenvelope, $E \equiv \infty$. Our key result is the Sex Entropy Theorem (5.5): If $\mathcal{H}$ is an entropy structure for (X, T), then E is a bounded affine superenvelope for the entropy sequence $\mathcal{H}$ if and only if there exists a symbolic extension φ of (X, T) such that $E = h^{\varphi}_{\text{ext}}$.

Consequently, we identify $h_{\text{sex}}(T, \cdot)$ as the infimum of all affine superenvelopes of $\mathcal{H}$, which equals $\mathrm{E}\mathcal{H}$, the smallest superenvelope of the entropy structure. We give two additional characterizations of this function, along with practical methods for estimation. The Entropy Structure Realization Theorem of [DS2] (see 8.4) shows that every nondecreasing bounded sequence of affine u.s.c. functions with u.s.c. differences on a Choquet simplex is realized as the entropy structure of a dynamical system. Along with the Sex Entropy Theorem, this reduces various questions of sex entropy to problems (or exercises) in functional analysis.

From here we describe the organization of the paper.

The work in the sections 2-4 is pure functional analysis. We use $\mathcal{F}$ to denote a nondecreasing sequence of nonnegative u.s.c. functions f_k with u.s.c. differences $f_{k+1} - f_k$ and $\lim f_k = f$. Entropy results follow by replacing $\mathcal{F}$ with $\mathcal{H}$. In Section 2, we introduce functional analytic background necessary for the sequel, and we characterize superenvelopes in terms of "continuous covers" of $\mathcal{F}$. In Section 3, we study the smallest superenvelope $\mathrm{E}\mathcal{F}$ of $\mathcal{F}$. We give an inductive characterization of $\mathrm{E}\mathcal{F}$ in terms of tail defects of upper semicontinuity along $\mathcal{F}$, using a construction which is often practical but in general can be unavoidably transfinite. We also provide estimates for $\mathrm{E}\mathcal{F}$. Work from this section is used in [DN], which provides the first C^r $(1 \leq r < \infty)$ examples of positive residual entropy. (If T is C^∞, then $h_{\text{res}}(T, \cdot) \equiv 0$ [BFF].) In Section 4, we give material on affine functions necessary to later entropy attainability results.

In Section 5, we give a careful statement of our key result, the Sex Entropy Theorem. In Section 6, we prove a zero dimensional case (the main argument) of this theorem. We explain the passage to the general case in Section 7. In Section 8, we assemble and discuss the applications of the preceding sections to symbolic extensions, including a variational principle for sex entropy and several attainability (and nonattainability) results. For the sex entropy and residual entropy functions, maximum values need not be attained on ergodic measures – a rare situation in ergodic theory. If $h_{\text{sex}}(T, \cdot)$ is affine (which is always the case when $K(X, T)$ is a Bauer simplex), then it will be the extension entropy function of a symbolic extension of T. On the other hand, it can happen that no symbolic extension of T has topological entropy equal to $\mathbf{h}_{\text{sex}}(T)$. We exhibit additional characterizations of asymptotic h-expansiveness, which is a very natural condition in the entropy structure/sex entropy context: T is asymptotically h-expansive if and only if it has a principal extension to a symbolic system (i.e., the sex entropy function h_{sex} equals h), if and only if h_k converges to h uniformly. We explain how the characterization in [D2] of the residual entropy of a system is a consequence of the superenvelope results. By mixing functional analytic examples of the earlier sections with the Entropy Structure Realization Theorem, we produce several examples of residual entropy phenomena. To be concrete, we end the section with two explicit constructions of dynamical systems: one with positive residual entropy, and another with zero residual entropy but with no symbolic extension of equal entropy.

Our entropy structures use refining sequences of partitions with small boundaries. To avoid partition constructions and the invocation of an auxiliary system in our general construction of an entropy structure (5.2(II)), and to get a better understanding of the dynamical meaning of sex entropy, it is appropriate to examine the relation of entropy to scale in terms of shrinking covers, continuous functions, epsilon-orbits with epsilon going to zero, and so on. Not all natural definitions give an entropy structure whose bounded affine superenvelopes are the sex entropy functions of symbolic extensions. The appropriate structures and equivalence relations are explained in [D3].

We thank the referees for their careful readings of the paper, and for their numerous useful suggestions which have improved it.

2. Semicontinuity, separation theorems and superenvelopes

This section contains certain general observations concerning the behavior of sequences of u.s.c. functions on compact metrizable spaces, including, in particular, affine u.s.c. functions on convex sets (simplexes). (For our applications we only consider the metrizable case, although most statements below hold without the metrizability assumption.)

Upper semicontinuous functions

We begin by recalling some basic facts about u.s.c. functions.

Definition 2.1. A function $f\colon K \mapsto \mathbb{R}$ defined on a topological Hausdorff space K is called *upper semicontinuous* (*u.s.c.*) if one of the following equivalent conditions holds.

(1) $f = \inf_\alpha g_\alpha$ for some family $\{g_\alpha\}$ of continuous functions.
(2) $f = \lim_\alpha g_\alpha$, where (g_α) is a nonincreasing net (sequence in metric spaces) of continuous functions.
(3) For each $r \in \mathbb{R}$, the set $\{x\colon f(x) \geq r\}$ is closed.
(4) $\limsup_{x'\to x} f(x') \leq f(x)$ at each $x \in K$.

Properties 2.2. (1) The infimum of any family of u.s.c. functions is again u.s.c. (by 2.1(1)).
(2) Both the sum and the supremum of finitely many u.s.c. functions are u.s.c. (by 2.1(3)).
(3) A uniform limit of u.s.c. functions is u.s.c.
(4) If $f = \lim f_\alpha$ is the limit of a nonincreasing net of u.s.c. functions defined on a compact domain K, and $g > f$ is a continuous function, then $g > f_\alpha$ for some α. In particular, a nonincreasing convergence of u.s.c. functions to a continuous limit is always uniform.
(5) Every u.s.c. function from a compact domain to $\mathbb{R}$ is bounded above and attains its maximum.

Notation 2.3. $\mathrm{SUP}(f)$ will be used to denote the supremum of f, i.e. $\sup_{x\in K} f(x)$, while $\sup_\alpha f_\alpha$ always means the function defined by pointwise supremum. We use INF and inf similarly.

Proposition 2.4 (**Exchanging Suprema and Infima).** *If* (f_α) *is a nonincreasing net of u.s.c. functions on a compact space, then* $\inf_\alpha \mathrm{SUP}(f_\alpha) = \mathrm{SUP}(\inf_\alpha f_\alpha)$.

Proof. The inequality "$\geq$" is obvious. To prove the converse, suppose that

$$\inf_\alpha \mathrm{SUP}(f_\alpha) > \mathrm{SUP}(\inf_\alpha f_\alpha)\ .$$

Viewing the left hand side as a constant function g we have $g > \inf_\alpha f_\alpha$. Thus, by Property 2.2(4), $g > f_\alpha(x)$ for some α and every point x, including the point where f achieves SUP. But $g \leq \mathrm{SUP}(f_\alpha)$ by definition, a contradiction. □

We now pass to lifting and projecting u.s.c. functions. Let $\pi\colon M \to K$ be a continuous map. If f is a continuous (or u.s.c.) function on K then the *lifted function* $f \circ \pi$ is a continuous (or u.s.c.) function on M. This lift we denote by the same letter f. We can also push functions down:

Definition 2.5. If $\pi\colon M \to K$ is a surjection and f is a bounded real-valued function on M, then $f^{[K]}$ is the function defined on K by

$$f^{[K]}(x) = \sup_{y \in \pi^{-1}(x)} f(y)\ .$$

Remark 2.6. In the case that f is u.s.c. and π is a continuous surjection between compact spaces, it is easily seen using Definition 2.1(3) that $f^{[K]}$ is also u.s.c.

Below, ex$K$ denotes the set of extreme points of a convex set K. A *Choquet simplex* is a compact convex subset K of a locally convex space such that the dual of the continuous affine functions on K is a lattice [AE, p.69]. We will only need the characterization of Choquet: a metrizable compact convex subset K of a locally convex space is a Choquet simplex if and only if for each x in K there exists a unique Borel probability μ_x supported on ex$K$ such that $\int f d\mu = f(x)$ for every continuous affine function on K [P1, p.60]. Every Choquet simplex is affinely homeomorphic to $K(X,T)$ for some homeomorphism T of a Cantor set [D1, O]; here the integral representation is the ergodic decomposition of a measure.

Proposition 2.7. *If* $\pi\colon M \to K$ *is a continuous affine surjection between convex sets, and* f *is an affine function defined on* M*, then* $f^{[K]}$ *is concave. If in addition* M *and* K *are Choquet simplexes and* π *preserves the extreme points (i.e.,* $\pi(\mathrm{ex}M) \subset \mathrm{ex}K$*), then* $f^{[K]}$ *is affine.*

Proof. Let $x = \alpha x_1 + \beta x_2$ in K ($\alpha \in (0,1),\ \beta = 1-\alpha$). Clearly

$$\pi^{-1}(x) \supset \{\alpha y_1 + \beta y_2\colon y_1 \in \pi^{-1}(x_1), y_2 \in \pi^{-1}(x_2)\}$$

hence concavity follows. For the reversed inclusion, and hence convexity of $f^{[K]}$ (with all the additional assumptions) we will apply the Choquet Representation Theorem and the Radon-Nikodym Theorem. Let μ_i be the unique probability measure supported by ex$K$ with barycenter at x_i, ($i = 1,2$). Clearly, $\mu = \alpha\mu_1 + \beta\mu_2$ is the unique probability measure supported by ex$K$ with barycenter at x, and each μ_i is absolutely continuous with respect to μ. Let f_i denote the corresponding Radon-Nikodym derivative defined on ex$K$. Note that $\alpha f_1 + \beta f_2 \equiv 1$. Let ν be the probability measure on ex$M$ with barycenter at a chosen point $y \in \pi^{-1}(x)$. Consider the measures ν_i on ex$M$ defined by $d\nu_i = (f_i \circ \pi)d\nu$ (here we use the assumption $\pi(\mathrm{ex}M) \subset \mathrm{ex}K$). We have

$$\nu_i(\mathrm{ex}M) = \langle f_i \circ \pi, \nu\rangle = \langle f_i, \pi(\nu)\rangle\ .$$

Since the barycenter of $\pi(\nu)$ coincides with the image by π of the barycenter of ν, i.e., with $\pi(y) = x$, by the uniqueness of such measure, we have $\pi(\nu) = \mu$. Thus

$$\nu_i(\mathrm{ex}M) = \langle f_i, \mu\rangle = 1,$$

which shows that each ν_i is a probability measure. Then we let y_i be the barycenter of ν_i. The image $\pi(y_i)$ coincides with the barycenter of the measure $\pi(\nu_i)$, and it is straightforward to verify that $\pi(\nu_i) = \mu_i$, hence $\pi(y_i) = x_i$ $(i = 1, 2)$. Finally,

$$\alpha d\nu_1 + \beta d\nu_2 = ((\alpha f_1 + \beta f_2) \circ \pi) d\nu = d\nu,$$

hence, passing to barycenters, $\alpha y_1 + \beta y_2 = y$, which completes the proof. □

Separation theorems

A function f is called *lower semicontinuous* (*l.s.c.*) if $-f$ is u.s.c. Essential to the sequel are various results, collected below, on separating l.s.c. and u.s.c. functions. We give some proof details where we lack precise references.

Theorem 2.8 (**Sandwich Theorem**). *Let K be a compact metric space. Suppose h and f are functions defined on a compact metric space K, $h \leq f$, h is u.s.c. and f is l.s.c. Then there exists a continuous function g such that $h \leq g \leq f$.*

Proof. See e.g. [T]. (The result holds for a topological space K if and only if K is normal.) □

Theorem 2.9 (**Separation of Disjoint Epigraphs**). *Suppose h and f are functions defined on a compact convex subset K of a locally convex linear space, $h < f$, h is u.s.c. and f is l.s.c.*

(1) *If h is concave and f is convex, then there exists an affine continuous function g such that $h < g < f$.*
(2) *If h is concave, then there exists a concave continuous function g such that $h < g < f$.*

Proof. (1) This is [C, Theorem 21.20], which we include for context.

(2) The concave u.s.c. function h is the pointwise infimum of the continuous affine functions a such that $h < a$ (see [C, Prop. 2.18]). Because f is l.s.c. with $h < f$, it follows that for each $x \in K$ there is an open neighborhood V and an affine function g_V such that $h < g_V$ and $g_V(y) < f(y)$ for $y \in V$. Let $g = \min g_V$, where the minimum is over a finite cover of K by such sets V. □

Theorem 2.10 (**Edwards' Separation Theorem**). *Suppose h and f are functions defined on a Choquet simplex K, $h \leq f$, h is convex and f is concave l.s.c. Then there exists an affine continuous function g such that $h \leq g \leq f$.*

Proof. See e.g. [AE], Theorem 7.6. □

Theorem 2.11. *Suppose h and f are functions defined on a Choquet simplex, $h < f$, h is affine u.s.c. and f is l.s.c. Then there exists an affine continuous function g such that $h < g < f$.*

Proof. Choose $\epsilon > 0$ such that $h + \epsilon < f$. By Theorem 2.9(2), there exists a concave continuous function g_1 such that $h + \epsilon < g_1 < f$. By Theorem 2.10, there exists an affine continuous function g_2 such that $h + \epsilon \leq g_2 \leq g_1$. Let $g = g_2$. □

Envelopes and superenvelopes

Let f be a bounded function defined on a compact domain K. We assume the set K is also convex in discussions involving convex structure. Throughout this paper we will be using the following notation:

Notation 2.12. $\widetilde{f}$ denotes the *u.s.c. envelope* of f, i.e., $\widetilde{f} = \inf_\alpha g_\alpha$, where g_α ranges over all continuous functions above f (i.e., $g_\alpha \geq f$).

In fact, $\widetilde{f}$ is the smallest u.s.c. function above f, and

$$\widetilde{f}(x) = \max\left\{ f(x), \quad \limsup_{x' \to x} f(x') \right\} .$$

It is also immediately seen that for any functions f and g, $\widetilde{f+g} \leq \widetilde{f} + \widetilde{g}$, with equality holding if f or g is continuous.

Definition 2.13. $\widehat{f}$ denotes the *u.s.c. concave envelope* of f, i.e., $\widehat{f} = \inf_\alpha g_\alpha$, where g_α range over all affine continuous functions above f.

Remarks 2.14. Clearly, $f \leq \widetilde{f} \leq \widehat{f}$. In fact, $\widehat{f}$ is the smallest concave u.s.c. function above f (see [C, Prop. 2.18]). If f is affine then $\widetilde{f}$ is concave, hence $\widetilde{f} = \widehat{f}$. Indeed, let $x = \alpha x_1 + \beta x_2$, and (using the above displayed formula for $\widetilde{f}$) choose sequences $x_{i,n} \to x_i$, $(i = 1, 2)$ such that $f(x_{i,n}) \to \widetilde{f}(x_i)$. Then

$$\widetilde{f}(x) \geq \lim_n f(\alpha x_{1,n} + \beta x_{2,n}) = \lim_n [\alpha f(x_{1,n}) + \beta f(x_{2,n})] = \alpha \tilde{f}(x_1) + \beta \tilde{f}(x_2).$$

Notation 2.15. If f is unbounded, we set $\widetilde{f} = \widehat{f} \equiv \infty$ which is to say that these envelopes do not exist (in the proper sense). Obviously, in any case $\mathrm{SUP}(f) = \mathrm{SUP}(\widetilde{f}) = \mathrm{SUP}(\widehat{f})$.

For the rest of this section, we set the following

Notation 2.16. $\mathcal{F}$ denotes $(f_k)_{k=1}^\infty$, a nondecreasing sequence of u.s.c. functions $f_k \colon K \to [0, +\infty)$, where K is compact metrizable, all differences $f_{k+1} - f_k$ are u.s.c., and $\lim_k f_k = f$. We allow ∞ as a possible value of f. We let f_0 denote the zero function.

Definition 2.17. *A superenvelope* of $\mathcal{F}$ is a function E on K such that either (i) E is bounded and for each $k \geq 1$ the function $E - f_k$ is nonnegative and u.s.c., or (ii) E is the constant function ∞.

Suppose E is a finite superenvelope of $\mathcal{F}$. Then $E = (E - f_1) + f_1$, so E is itself a u.s.c. function. The functions $E - f_k$ converge nonincreasingly to the function $E - f$, which hence is also u.s.c. and nonnegative. Since the pointwise minimum of two superenvelopes is easily seen to be again a superenvelope, the superenvelopes of $\mathcal{F}$ form a directed family. It follows immediately from Property 2.2(1) that there exists a smallest superenvelope of $\mathcal{F}$, namely, the pointwise infimum (limit) of all superenvelopes. We denote that superenvelope by $\mathrm{E}\mathcal{F}$. This function is either bounded or it is the constant ∞.

We will now provide our first construction of the smallest superenvelope $\mathrm{E}\mathcal{F}$. By a *continuous cover* (or simply *cover*) of $\mathcal{F}$ we shall mean any sequence $\mathcal{G} = (g_k)_{k \in \mathbb{N}}$ of continuous functions from K to $[0, \infty]$ such that $g_k \geq f_k - f_{k-1}$ for each $k \in \mathbb{N}$. Let $\Sigma\mathcal{G}$ denote $\Sigma_{k=1}^\infty g_k$. We will consider the u.s.c. function

$$\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}},$$

where the infimum is taken over all covers $\mathcal{G}$ of $\mathcal{F}$.

Theorem 2.18. $\mathrm{E}\mathcal{F} = \inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}}$. *There exists a cover* $\mathcal{G}$ *with* $\mathrm{SUP}(\Sigma\mathcal{G}) = \mathrm{SUP}(\mathrm{E}\mathcal{F})$.

Proof. Notice that the set $\{\mathcal{G}\}$ of all covers of $\mathcal{F}$ is a directed family, with the order $\mathcal{G} \succcurlyeq \mathcal{G}'$ if $g_k \le g'_k$ for each k; the common successor of two covers is obtained as their pointwise minimum. By attaching to each cover $\mathcal{G}$ the function $\widetilde{\Sigma\mathcal{G}}$, we regard $\{\widetilde{\Sigma\mathcal{G}}\}$ as a net indexed by $\mathcal{G}$. This is a nonincreasing net of u.s.c. functions, and the infimum in the definition of $\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}}$ is in fact a nonincreasing limit.

Let $\mathcal{F}_k$ denote the sequence $(f_{n+k} - f_k)_{n\in\mathbb{N}}$. We can write

$$\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}} = \lim_{\mathcal{G}} \left(\Sigma_{i=1}^k g_i + \widetilde{\Sigma\mathcal{G}_k} \right) = f_k + \inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}_k} \ ,$$

where $\mathcal{G}_k = (g_{n+k})_{n\in\mathbb{N}}$ is the cover of $\mathcal{F}_k$ obtained by "truncating" $\mathcal{G}$ (in fact all covers of $\mathcal{F}_k$ are such truncations). The difference between the function $\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}}$ and f_k is hence a nonnegative u.s.c. function. We have proved that the function $\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}}$ is a superenvelope of $\mathcal{F}$, hence it is not smaller than $\mathrm{E}\mathcal{F}$.

The converse inequality holds trivially if $\mathrm{E}\mathcal{F} \equiv \infty$. Suppose $\mathrm{E}\mathcal{F}$ is bounded and let $g \ge \mathrm{E}\mathcal{F}$ be a continuous function. We will inductively construct a cover $\mathcal{G}$ such that $\Sigma\mathcal{G} \le g$. This will end the proof, because then also $\widetilde{\Sigma\mathcal{G}} \le g$ and hence

$$\inf_{\mathcal{G}} \widetilde{\Sigma\mathcal{G}} \le \inf_{g \ge \mathrm{E}\mathcal{F}} g = \mathrm{E}\mathcal{F}.$$

(1) It is clear that

$$f_1 \le g - (\mathrm{E}\mathcal{F} - f_1)$$

and in this inequality the assumptions of the Sandwich Theorem 2.8 hold. Thus, there exists a continuous function g_1 in between.

(2) Suppose we have found continuous functions $g_1, g_2, \dots, g_k$ such that we have $g_i \ge f_i - f_{i-1}$ for $i = 1, 2, \dots, k$, and

$$\sum_{i=1}^k g_i \le g - (\mathrm{E}\mathcal{F} - f_k) \ .$$

(This condition was fulfilled in (1) above for the case $k = 1$.) Then

$$f_{k+1} - f_k \le g - \sum_{i=1}^k g_i - (\mathrm{E}\mathcal{F} - f_{k+1})$$

and again, the assumptions of the Sandwich Theorem are being fulfilled. Thus a continuous function g_{k+1} exists between the expressions on both sides. It is seen that the inductive assumption is now satisfied for $k+1$. In this manner a cover $\mathcal{G} = (g_k)_{k\in\mathbb{N}}$ of $\mathcal{F}$ is constructed. Finally, since each $\mathrm{E}\mathcal{F} - f_k$ is nonnegative, $\Sigma_{i=1}^k g_i \le g$ holds for each k, hence $\Sigma\mathcal{G} \le g$.

The cover mentioned in the last statement of the theorem is obtained by applying the last argument to the constant function $g \equiv \mathrm{SUP}(\mathrm{E}\mathcal{F})$. □

Example 2.19. Let K consist of a sequence $(a_n)_{n\ge 2}$ and its limit b. Define $f_1 = 1_{\{b\}}$ and $f_k = f_{k-1} + 1_{\{a_k\}}$ if $k \ge 2$. Clearly $f \equiv 1$, and it is not hard to see using continuous covers (or Proposition 3.13) that $\mathrm{E}\mathcal{F} = 1_{\{b\}} + 1$. In this example, it is seen that f being u.s.c. (even continuous) does not imply that it is equal to $\mathrm{E}\mathcal{F}$.

Example 2.20. Let K be a space without isolated points, and let $(q_n)_{n\in\mathbb{N}}$ be a dense sequence. Set $f_1 = 1_{q_1}$ and $f_k = f_{k-1} + 1_{\{q_k\}}$ if $k \geq 2$. Clearly $f \leq 1$; however, $\mathrm{E}\mathcal{F} \equiv \infty$. To see this, check that $\Sigma\mathcal{G}$ is unbounded for any continuous cover $\mathcal{G}$, or apply Proposition 3.9.

Example 2.21. Let K be the following subset of $\mathbb{R}^2$:

$$K = \{(0,0)\} \cup \{(\tfrac{1}{m},0)\colon\ m \in \mathbb{N}\} \cup \{(\tfrac{1}{m},\tfrac{1}{n})\colon (m,n) \in \mathbb{N}^2,\ m \leq n\}\ .$$

For $k \in \mathbb{N}$ and given $a,b,c \in \mathbb{R}$, define $f_k = \sum_{i=1}^{k} d_i$, where

$$d_1 = a \cdot 1_{\{(0,0)\}}\ , \qquad d_{2k} = b \cdot 1_{\{(\frac{1}{k},0)\}}\ , \qquad d_{2k+1} = c \cdot 1_{\{(\frac{1}{m_k},\frac{1}{n_k})\}}\ ,$$

and (m_k,n_k) is some linear ordering of the set $\{(m,n) \in \mathbb{N}^2\colon\ m \leq n\}$. Clearly $\mathrm{SUP}(f) = \max\{a,b,c\}$. By analyzing continuous covers (or computing $\mathrm{E}\mathcal{F}$ from Theorem 3.3), one can check $\mathrm{E}\mathcal{F}\left((\frac{1}{m},\frac{1}{n})\right) = c$, $\mathrm{E}\mathcal{F}\left((\frac{1}{m},0)\right) = b{+}c$, and $\mathrm{E}\mathcal{F}\left((0,0)\right) = a+b+c$. Playing with higher order accumulation points in K, we can obtain arbitrary real numbers $0 < r_1 \leq r_2$ as $\mathrm{SUP}(f)$ and $\mathrm{SUP}(\mathrm{E}\mathcal{F})$.

3. Estimates and inductive characterization of $\mathrm{E}\mathcal{F}$

Throughout this section we continue the notation and assumptions of (2.16), and we also assume that f is bounded (otherwise there is no point in estimating $\mathrm{E}\mathcal{F}$). First of all, it is important to realize that when $\mathrm{E}\mathcal{F}$ is bounded, the functions $\mathrm{E}\mathcal{F}$ and f must be equal on a large set.

Proposition 3.1. *If* $\mathrm{E}\mathcal{F} < \infty$, *then* $\mathrm{E}\mathcal{F} = f$ *on a residual subset of* K.

Proof. Suppose $\mathrm{E}\mathcal{F} - f \geq \epsilon$ on a set containing an open set U. Then $\mathrm{E}\mathcal{F} - \epsilon 1_U$ is easily seen to be a superenvelope of $\mathcal{F}$, which contradicts minimality of $\mathrm{E}\mathcal{F}$. We have shown that the set where $\mathrm{E}\mathcal{F} - f \geq \epsilon$ has empty interior. This set is also closed, so $\mathrm{E}\mathcal{F} > f$ on a first category set. □

On the other hand, for f to equal $\mathrm{E}\mathcal{F}$ exactly is rather special.

Proposition 3.2. *The following conditions are equivalent:*

(1) $\mathrm{E}\mathcal{F} = f$
(2) f_k *converges to* f *uniformly on* K.

Proof. If (1) holds, then $f - f_k$ is a sequence of u.s.c. functions which converge nonincreasingly to zero. By Property 2.2(4), the convergence is uniform, so (2) holds.

If (2) holds, then every $f - f_k$ is a uniform limit of u.s.c. functions, hence is u.s.c., so f is a superenvelope, and obviously it is the minimal one, so (1) holds. □

Evidently the mystery of $\mathrm{E}\mathcal{F} - f$ has to do with upper semicontinuity and nonuniform convergence of the sequence $\mathcal{F}$. We will capture this with an inductive characterization, which is often practical but in general transfinite. Given a nonincreasing sequence $\mathcal{T} = (t_k)$ of nonnegative bounded functions on K, we set $u_0 \equiv 0$, and then for an ordinal α set

$$u_{\alpha+1} = \lim_{k\to\infty} \widetilde{u_\alpha + t_k}\ .$$

Notice that if u_α is bounded, then $\widetilde{u_\alpha + t_k}$ converges nonincreasingly, and $u_{\alpha+1}$ is also nonnegative u.s.c. and bounded. Finally, for a limit ordinal β let

$$u_\beta = \widetilde{\sup_{\alpha<\beta} u_\alpha}\ .$$

(So, for every α, either u_α is bounded or by our convention (2.15) $u_\alpha \equiv \infty$.) We will apply this construction to the tail sequence $\mathcal{T} = (\tau_k)_k$ of the sequence $\mathcal{F}$, defined by $\tau_k = f - f_k$ for $k \geq 0$. For this choice of $\mathcal{T}$ we may write u_α as $u_\alpha^{\mathcal{F}}$.

Theorem 3.3 (**Inductive Characterization).** *Let $\mathcal{F}$ be a nondecreasing sequence of nonnegative u.s.c. functions converging to a bounded limit f, and with u.s.c. differences $f_{k+1} - f_k$. Let $\mathcal{T} = (\tau_k)_k$ denote the sequence of tails. The functions $u_\alpha = u_\alpha^{\mathcal{F}}$ are nondecreasing in α, and*

$$u_\alpha = u_{\alpha+1} \iff u_\alpha = \mathrm{E}\mathcal{F} - f. \tag{3.4}$$

Moreover, such an α exists already among countable ordinals. Consequently,

$$\mathrm{E}\mathcal{F} - f = \sup_\alpha u_\alpha \tag{3.5}$$

where the supremum is taken over all countable ordinals α, or equivalently over all ordinals α.

Proof. That the functions u_α are nondecreasing in α is obvious from the definition. We will first show

$$u_\alpha \leq \mathrm{E}\mathcal{F} - f \tag{3.6}$$

for any α. This is clear for $\alpha = 0$, so suppose it holds for some α. Write

$$\mathrm{E}\mathcal{F} - f_k = \mathrm{E}\mathcal{F} - f + \tau_k \geq u_\alpha + \tau_k.$$

Because the function on the left is u.s.c., we can replace the function on the right with its u.s.c. envelope, so

$$\mathrm{E}\mathcal{F} - f_k \geq \widetilde{u_\alpha + \tau_k} \geq u_{\alpha+1}.$$

Passing with k to infinity we obtain $u_{\alpha+1} \leq \mathrm{E}\mathcal{F} - f$. If β is a limit ordinal and (3.6) holds for $\alpha < \beta$, then the inequality $u_\beta \leq \mathrm{E}\mathcal{F} - f_k$ follows from the definition of u_β and the uppersemicontinuity of $\mathrm{E}\mathcal{F} - f$. This proves (3.6) for all α.

We will now prove (3.4). If $u_\alpha \equiv \infty$, then by (3.6) the right side of (3.4) holds, and obviously the left side holds. Suppose then u_α is bounded. If the right side of (3.4) holds, then

$$u_\alpha + \tau_k = \mathrm{E}\mathcal{F} - f + \tau_k = \mathrm{E}\mathcal{F} - f_k$$

which is u.s.c., so we also have $\widetilde{u_\alpha + \tau_k} = \mathrm{E}\mathcal{F} - f_k$. Thus,

$$u_{\alpha+1} = \lim_k (\mathrm{E}\mathcal{F} - f_k) = \mathrm{E}\mathcal{F} - f = u_\alpha.$$

Now assume the left side of (3.4) holds. It suffices to bound, for each k, the "defect of upper semi-continuity" of $u_\alpha + \tau_k$ as follows:

$$(\widetilde{u_\alpha + \tau_k}) - (u_\alpha + \tau_k) \leq u_{\alpha+1} - u_\alpha.$$

Once this is proved, $u_{\alpha+1} = u_\alpha$ implies that $u_\alpha + \tau_k$ is u.s.c. for every k; then $u_\alpha + \tau_k = u_\alpha + f - f_k$ shows $u_\alpha + f$ is a superenvelope; then $u_\alpha + f \geq \mathrm{E}\mathcal{F}$, which together with (3.6) gives equality. To estimate these defects first note that they converge to $u_{\alpha+1} - u_\alpha$ (because τ_k converges to zero). It now suffices to show that they are nondecreasing in k. Let $\ell > k$. We have

$$\begin{aligned}
&(\widetilde{u_\alpha + \tau_\ell}) - (u_\alpha + \tau_\ell) - (\widetilde{u_\alpha + \tau_k}) + (u_\alpha + \tau_k) \\
&= (\widetilde{u_\alpha + \tau_\ell}) - \ u_\alpha - \tau_\ell \ - (u_\alpha + \tau_\ell + f_\ell - f_k)^{\sim} + u_\alpha + \tau_\ell + f_\ell - f_k \\
&= (\widetilde{u_\alpha + \tau_\ell}) + \ f_\ell - f_k \ - (u_\alpha + \tau_\ell + f_\ell - f_k)^{\sim} \ \geq \ 0
\end{aligned}$$

because $f_\ell - f_k$ is u.s.c. This completes the proof of the equivalence statement.

For the "Moreover" claim, let $K_\alpha = \{(x,y) \in K\times[0,\infty] \colon 0 \le y \le u_\alpha(x)\}$, a compact set. The space of compact subsets of $K \times [0,\infty]$ is separable in the Hausdorff metric, and $\alpha < \beta < \gamma$ implies $\mathrm{dist}(K_\alpha, K_\beta) \le \mathrm{dist}(K_\alpha, K_\gamma)$, by inclusion. Therefore, given $\epsilon > 0$ there can be only countably many α such that $\mathrm{dist}(K_\alpha, K_{\alpha+1}) > \epsilon$. It follows that $u_\alpha = u_{\alpha+1}$ for some countable ordinal α.

The final claim (3.5) is now obvious. □

Examples (which we skip) show that the transfiniteness in the induction is necessary; on the other hand, sometimes one can compute $\mathrm{E}\mathcal{F}$ by computing u_α and $u_\alpha = u_{\alpha+1}$ for a finite α (e.g., $u_1 \ne u_2 = u_3$ in Example 2.21). The rest of this section will be devoted to some practical estimates for $\mathrm{E}\mathcal{F}$. We will use the especially accessible function $u := u_1^{\mathcal{F}} = \inf_k \widetilde{\tau}_k = \lim_k \widetilde{\tau}_k$.

Proposition 3.7. *Let τ_k and $u_1^{\mathcal{F}}$ be defined as in Theorem 3.3, and set $u = u_1^{\mathcal{F}}$.*

1. *The functions $\widetilde{\tau}_k - \tau_k$ converge nondecreasingly in k to the limit u.*
2. *$\tilde{f} - f \le u \le \tilde{f}$.*
3. *The pointwise supremum of u is $\inf_k \mathrm{SUP}(\tau_k) = \lim_k \mathrm{SUP}(\tau_k)$.*

Proof. (1) The proof of Theorem 3.3 showed the defects $(\widetilde{u_\alpha + \tau_k}) - (u_\alpha + \tau_k)$ converge nondecreasingly in k to $u_{\alpha+1} - u_\alpha$. The case $\alpha = 0$ is (1).

(2) The first inequality follows from (1) because $\tilde{f} - f = \widetilde{\tau}_0 - \tau_0$, and the second inequality is trivial by the definition of $u = u_1$.

(3) By the Exchanging Suprema and Infima statement of Proposition 2.4,

$$\inf_k \mathrm{SUP}(\tau_k) = \inf_k \mathrm{SUP}(\widetilde{\tau}_k) = \mathrm{SUP}(\inf_k \widetilde{\tau}_k) = \mathrm{SUP}(u).$$

□

We recall a standard definition from topology

Definition 3.8. A point x has *order of accumulation* 0 in a set K when it is an isolated point in K, i.e., when $\{x\}$ is open. Let $K^{(1)}$ denote the compact set of all points that are not of order 0. Inductively, x is (an accumulation point) *of order* $r = r(x)$ $(r \ge 1)$ in K if it is an isolated point of the compact set $K^{(r)}$ of points which are not of order smaller than r. We set $r(x) = \infty$ for points which are not of any finite order.

Proposition 3.9. *For $t > 0$, let $r(t,x)$ denote the order of accumulation of x in the set K_t of points where $u \ge t$. Then $\mathrm{E}\mathcal{F}(x) \ge f(x) + tr(t,x)$.*

Proof. By the characterization (3.5) of $\mathrm{E}\mathcal{F} - f$ in Theorem 3.3, it suffices to prove the claim: if $r(t,x) \ge m$, then $u_m(x) \ge mt$. The case $m = 0$ is trivial. Suppose the claim holds for m, and $r(t,x) \ge m+1$. Choose points x_n from K_t such that $\lim x_n = x$ and $r(x_n) \ge m$. Then $u_{m+1}(x) \ge \overline{\lim}_n u_{m+1}(x_n) = \overline{\lim}_n \lim_k (u_m(x_n) + \tau_k(x_n))^\sim \ge \overline{\lim}_n \lim_k (mt + \tau_k(x_n))^\sim \ge mt + \overline{\lim}_n u(x_n) \ge (m+1)t$. □

From here we concern ourselves with bounds on $\mathrm{E}\mathcal{F}$ from above.

Proposition 3.10. *Let $r(x)$ denote the order of accumulation of x in the domain K of f. We have the following estimate of the minimal superenvelope $\mathrm{E}\mathcal{F}$:*

$$\mathrm{E}\mathcal{F}(x) \le f(x) + r(x)u(x).$$

Proof. We proceed by induction on r. By Proposition 3.1, $\mathrm{E}\mathcal{F} = f$ at any isolated point, so the statement holds for $r = 0$. Suppose we have proved it for some $r \geq 0$. Let x be a point of order $r + 1$. Define E as $\mathrm{E}\mathcal{F}$ except at x where we set $E(x) = f(x) + (r + 1)u(x)$. We complete the proof by showing that E is a superenvelope of $\mathcal{F}$. The functions $E - f_k$ are nonnegative. We only need to verify their upper semi-continuity at x, i.e., given k we need

$$\limsup_{x' \to x}(E - f_k)(x') \leq (E - f_k)(x).$$

The point x is surrounded only by points $x' \neq x$ of order at most r, for which

$$\begin{aligned}(E - f_k)(x') - (E - f_k)(x) &= \mathrm{E}\mathcal{F}(x') - f_k(x') - f(x) - (r+1)u(x) + f_k(x) \\ &\leq f(x') + ru(x') - f_k(x') - f(x) - (r+1)u(x) + f_k(x) \\ &= \tau_k(x') - \tau_k(x) + r\big(u(x') - u(x)\big) - u(x) \ .\end{aligned}$$

Passing to limsup over $x' \to x$, we appeal to upper semicontinuity of u and then Proposition 3.7(1) to obtain

$$\limsup_{x' \to x}(E - f_k)(x') - (E - f_k)(x) \ \leq \ \widetilde{\tau}_k(x) - \tau_k(x) - u(x) \ \leq \ 0$$

as required. □

A refinement of the above method leads to another estimate:

Theorem 3.11. *For $t > 0$ and $x \in X$, define $r_t(x)$ to be 0 if $u(x) < t$, and otherwise define $r_t(x)$ to be 1 plus the order of accumulation of x in the set $\{x\colon u(x) \geq t\}$ (i.e., $r_t(x) = 1 + r(t, x)$). Suppose there is an open set V and a function $g\colon \mathbb{R} \to \mathbb{R}$ such that $r_t(x) \leq g(t)$ for all $x \in V$, and $\int_0^\infty g(t)dt < \infty$. Define the nonnegative function*

$$E(x) = f(x) + \int_0^\infty r_t(x)dt \ .$$

Then $E - f_k$ is u.s.c. on V. If $\mathrm{E}\mathcal{F}$ is bounded, then $\mathrm{E}\mathcal{F}(x) \leq E(x)$ for all $x \in V$.

Proof. The integral is unaffected if the upper limit ∞ is changed to $u(x)$. Let x_n be a sequence from $V \setminus \{x\}$ with $\lim x_n = x$. We need to show for $k \in \mathbb{N}$ that

$$\limsup_n (E - f_k)(x_n) \leq (E - f_k)(x).$$

Given t with $0 < t \leq u(x)$, all points sufficiently close to x have order of accumulation in $\{x\colon u(x) \geq t\}$ smaller than that of x, i.e., $\limsup_n r_t(x_n) \leq r_t(x) - 1$. For $t > u(x)$, by upper semicontinuity of u, $\limsup_n r_t(x_n) = 0$. Thus, for all $t > 0$ we can write $\limsup_n r_t(x_n) \leq \max\{r_t(x) - 1, 0\}$. By the Dominated Convergence Theorem,

$$\limsup_n \int_0^\infty r_t(x_n)dt \leq \int_0^{u(x)} r_t(x)dt - u(x)$$

and therefore (with $\tau_k = f - f_k$)

$$\begin{aligned}\limsup_n (E - f_k)(x_n) \ &\leq \ \limsup_n \tau_k(x_n) + \limsup_n \int_0^\infty r_t(x_n)dt \\ &\leq \widetilde{\tau}_k(x) + \int_0^{u(x)} r_t(x)dt - u(x) \\ &\leq \ \widetilde{\tau}_k(x) + \int_0^\infty r_t(x)dt - (\widetilde{\tau}_k(x) - \tau_k(x))\end{aligned}$$

because u estimates from above the defect of τ_k. The last expression, after cancellation, coincides with $(E - f_k)(x)$.

Now suppose that $\mathrm{E}\mathcal{F}$ is bounded. If M is a constant such that $E \le M$ on V (e.g. $M = \mathrm{SUP}(f) + \int_0^\infty g(t)dt$), then the function which equals E on V and equals $\mathrm{E}\mathcal{F} + M$ outside V is a superenvelope. This proves the final claim. □

Remark 3.12. The domination assumption cannot be dropped. For instance, in general $u(x) = 0$ does not imply $\mathrm{E}\mathcal{F}(x) = f(x)$ (it does whenever Theorem 3.11 applies). For example, x can be the limit of closed sets C_n with $\mathrm{SUP}(u|C_n) = \frac{1}{n}$, $u(x) = f(x) = 0$, but $\mathrm{SUP}(\mathrm{E}\mathcal{F}|C_n) = 1$, hence $\mathrm{E}\mathcal{F}(x) = 1$.

Proposition 3.13. *Suppose j is a nonnegative integer such that $f_{k+1} - f_k$ is continuous whenever $k \ge j$. Then $\mathrm{E}\mathcal{F} = f_j + \widetilde{\tau_j}$, and $\mathrm{SUP}(\mathrm{E}\mathcal{F}) \le 2\mathrm{SUP}(f)$.*
If all f_k are continuous, then $\mathrm{E}\mathcal{F} = \widetilde{f}$ and $\mathrm{SUP}(\mathrm{E}\mathcal{F}) = \mathrm{SUP}(f)$.

Proof. Let E denote the function $f_j + \widetilde{\tau_j}$. Since $\mathrm{E}\mathcal{F} - f_j$ is u.s.c. and majorizes τ_j, it also majorizes $\widetilde{\tau_j} = E - f_j$, hence $\mathrm{E}\mathcal{F} \ge E$. On the other hand, for every k the difference $f_j - f_k$ is u.s.c. (consider the cases $k > j$ and $k \le j$), so

$$\begin{aligned} E - f_k &= f_j + \widetilde{\tau_j} - f_k \quad \text{ is u.s.c. and} \\ E - f_k &\ge f_j + \tau_j - f_k = f - f_k \ge 0. \end{aligned}$$

Therefore E is a superenvelope, and $E \ge \mathrm{E}\mathcal{F}$. The remaining statements of the assertion are now obvious. □

4. Affine functions and attainability

In this section we continue to assume the conditions 2.16 governing $\mathcal{F}$ and in addition we suppose the common domain K of the functions f_k is a Choquet simplex and the functions f_k are affine. We say a cover $\mathcal{G}$ of $\mathcal{F}$ is *offset* if $g_k > f_k - f_{k-1}$ for every k.

Proposition 4.1. *Given $\mathcal{F}$ and K as above, the following hold.*

$$\mathrm{E}\mathcal{F} = \inf_{\mathcal{G}_\mathrm{A}} \widetilde{\widehat{\Sigma\mathcal{G}}}_\mathrm{A} = \inf_{\mathcal{G}_\mathrm{A}} \widehat{\Sigma\mathcal{G}}_\mathrm{A} = \inf_{\mathcal{G}'_\mathrm{A}} \widehat{\Sigma\mathcal{G}'}_\mathrm{A} = \lim_{\mathcal{G}'_\mathrm{A}} \widehat{\Sigma\mathcal{G}'}_\mathrm{A} \tag{4.2}$$

where $\mathcal{G}_\mathrm{A}$ ranges over all affine covers of $\mathcal{F}$, and $\mathcal{G}'_\mathrm{A}$ ranges over all affine offset covers of $\mathcal{F}$. The u.s.c. function $\mathrm{E}\mathcal{F}$ is concave.

Proof. Obviously, in the formula $\mathrm{E}\mathcal{F} = \inf_\mathcal{G} \widetilde{\Sigma\mathcal{G}}$ of Theorem 2.18, we can take the infimum only over offset covers. Then, using Theorem 2.11, every such cover dominates another cover $\mathcal{G}_\mathrm{A}$ consisting of affine functions. This proves the first equality. The second equality holds, as explained in Remarks 2.14, because the functions $\Sigma\mathcal{G}_\mathrm{A}$ are affine. (This shows $\mathrm{E}\mathcal{F}$ is concave.) The next equality is obvious and the last follows because, by Theorem 2.11 again, the affine offset covers form a directed family. □

Another function of interest to us in this case is

$$\mathrm{E}_\mathrm{A}\mathcal{F} = \inf E_\mathrm{A},$$

defined as the infimum of all affine superenvelopes of $\mathcal{F}$. This is also a concave nonnegative u.s.c. function. In fact we show the following equalities:

Theorem 4.3.
$$\mathrm{E_A}\mathcal{F} = \mathrm{E}\mathcal{F}, \quad \inf_{E_\mathrm{A}} \mathrm{SUP}(E_\mathrm{A}) = \mathrm{SUP}(\mathrm{E}\mathcal{F}).$$

Remark 4.4. Note that since the affine superenvelopes usually DO NOT form a directed family (see Example 4.7), the second equality is not a consequence of the first and a statement on exchanging suprema and infima as in Proposition 2.4.

Proof of Theorem 4.3. The inequalities "$\geq$" are obvious and equalities hold if $\mathrm{E}\mathcal{F} \equiv \infty$. Suppose $\mathrm{E}\mathcal{F}$ is bounded and let g be an affine continuous function above the sum of some fixed affine cover, $g \geq \Sigma\mathcal{G}_\mathrm{A}$, $\mathcal{G}_A = (g_k)_{k=1}^\infty$. We will find an affine superenvelope E_A below g. Set

$$E_\mathrm{A} = g - (\Sigma\mathcal{G}_\mathrm{A} - f)$$

This E_A is an affine function not larger than g. Moreover, for each $k \in \mathbb{N}$,

$$E_\mathrm{A} - f_k = (g - \Sigma\mathcal{G}_\mathrm{A}) + \sum_{n=k+1}^{\infty} (f_n - f_{n-1}) = g - \sum_{n=1}^{k} g_n - \sum_{n=k+1}^{\infty} (g_n - (f_n - f_{n-1})),$$

which is nonnegative (see middle expression) and u.s.c. (each $g_n - (f_n - f_{n-1})$ is l.s.c., and so is the sum as a nondecreasing limit). So, E_A is an affine superenvelope of $\mathcal{F}$. Thus $\mathrm{E_A}\mathcal{F} \leq \widehat{\Sigma\mathcal{G}_\mathrm{A}}$. But this is true for any $\Sigma\mathcal{G}_\mathrm{A}$, so, by (4.2), $\mathrm{E_A}\mathcal{F} \leq \mathrm{E}\mathcal{F}$. For the second statement let $g \equiv \mathrm{SUP}(\mathrm{E}\mathcal{F}) + \epsilon$. Using Property 2.2(4) and the fact that affine offset covers are directed, we can find an affine offset cover $\mathcal{G}'_\mathrm{A}$ with $\widetilde{\Sigma\mathcal{G}'_\mathrm{A}} < g$. The above argument now produces an affine superenvelope E_A below $\mathrm{SUP}(\mathrm{E}\mathcal{F}) + \epsilon$, which ends the proof. □

Given x in a metrizable Choquet simplex K, let μ_x be the unique Borel probability on ex$K$ such that for every affine continuous (or u.s.c.) function a on K,

$$a(x) = \int_{\mathrm{ex}K} (a|_{\mathrm{ex}K}) d\mu_x$$

where $a|_{\mathrm{ex}K}$ is the restriction of a to ex$K$. The point x is called the barycenter of the measure μ_x. The restriction map establishes a 1-1 order preserving correspondence between affine continuous functions on K and certain continuous functions on ex$K$ (namely those which are restrictions of affine continuous functions on K). We can define

$$g^{\mathrm{aff}}(x) = \int_{\mathrm{ex}K} g d\mu_x$$

for any bounded continuous (or u.s.c.) function g on ex$K$, but in general this function need not be continuous (u.s.c.), because the map $x \mapsto \mu_x$ need not be continuous.

A *Bauer simplex* is a Choquet simplex K such that the extreme set ex$K$ is compact. In this case the map $x \mapsto \mu_x$ is a continuous affine bijection onto the Borel probabilities on ex$K$ [P1, pp.65-6]. Consequently, for a Bauer simplex K the map $f \mapsto f|_{\mathrm{ex}K}$ gives a 1-1 correspondence between affine continuous functions on K and all continuous functions on ex$K$, and the same formula establishes a 1-1 correspondence between all nonnegative affine u.s.c. functions on K and all nonnegative u.s.c. functions on ex$K$. Consequently, for a Choquet simplex K, if g is nonnegative continuous (u.s.c.) on ex$K$ and vanishes outside a compact subset C of ex$K$, then g^{aff} is necessarily nonnegative continuous (u.s.c.) on K (because the closed convex hull of C is a Bauer simplex with extreme set C, and elsewhere on K the function g^{aff} is 0).

The discussion of Bauer simplicies justifies the easy

Remark 4.5. Our examples 2.19, 2.20, 2.21 can be adapted to produce corresponding affine examples on Bauer simplices, by viewing the given compact domain as the extreme set of a simplex and extending all functions affinely.

If now $\mathcal{F}$ is a sequence of affine nonnegative u.s.c. functions on a Bauer simplex K then the restriction map is a 1-1 correspondence between all affine superenvelopes of $\mathcal{F}$ and all superenvelopes of $\mathcal{F}|_{\text{ex}K}$. But this implies that the affine superenvelopes of $\mathcal{F}$ form a directed family and that the smallest affine superenvelope of $\mathcal{F}$ exists. Combining this with Theorem 4.3 we conclude:

Theorem 4.6. *If $\mathcal{F}$ is a sequence of nonnegative affine u.s.c. functions defined on a Bauer simplex K, then $\mathrm{E}\mathcal{F}$ is affine.*

In the following example we show that for general Choquet simplexes the situation may be significantly different: the smallest affine superenvelope need not exist, and the infimum of pointwise suprema of such envelopes may never be attained. (This example also shows that affine superenvelopes need not form a directed family, because the limit would equal the smallest superenvelope and it would be affine.)

Example 4.7. We thank Bob Phelps [P2] for showing us the road that leads to this example. It is an example of a sequence $\mathcal{F}$ of nonnegative affine u.s.c. functions defined on a Choquet simplex K for which $\mathrm{SUP}(E_{\mathrm{A}}) > \mathrm{SUP}(\mathrm{E}\mathcal{F})$ $(= \inf_{E_{\mathrm{A}}} \mathrm{SUP}(E_{\mathrm{A}})$, by the second statement of Theorem 4.3), for every affine superenvelope E_{A} of $\mathcal{F}$. Moreover, in the example $\mathrm{E}\mathcal{F} = \widetilde{f}$, and in particular $\mathrm{SUP}(\mathrm{E}\mathcal{F}) = \mathrm{SUP}(f)$.

Let K be the Choquet simplex whose set of extreme points consists of a point b_1, a sequence $(a_n)_{n\ge 1}$ converging to b_1, and a sequence $(b_n)_{n\ge 2}$ converging to

$$b = \sum_{n=1}^{\infty} 2^{-n} a_n.$$

Define $\mathcal{F} = (f_k)_{k\in\mathbb{N}}$, where $f_k = (\sum_{n=1}^{k} 1_{\{b_n\}})^{\text{aff}}$. Since characteristic functions of single points have compact supports, f_1 is an affine u.s.c. function, and $f_k - f_{k-1}$ is an affine continuous function if $k \ge 2$. Consequently by Proposition 3.13,

$$\mathrm{E}\mathcal{F} = (1_{\{b_1\}})^{\text{aff}} + (\sum_{n=2}^{\infty} (1_{\{b_n\}})^{\text{aff}})^{\sim} \ .$$

By Theorem 4.9(1) below, $\mathrm{E}\mathcal{F}$ must assume its maximum on the closure of the extreme points. The maximum of $\mathrm{E}\mathcal{F}$ on $\text{ex}K$ is 1. The only new accumulation point of $\text{ex}K$ is b, and since $(1_{\{b_1\}})^{\text{aff}}(b) = 0$ we have

$$\mathrm{E}\mathcal{F}(b) = 0 + (\sum_{n=2}^{\infty} (1_{\{b_n\}})^{\text{aff}})^{\sim}(b) \le 0 + 1 \ .$$

Consequently $\mathrm{SUP}(\mathrm{E}\mathcal{F}) = 1$.

Now suppose E_{A} is an affine superenvelope of $\mathcal{F}$ not exceeding 1. Clearly, $E_{\mathrm{A}}(b_n) = 1$ for each n. By upper semi-continuity, $E_{\mathrm{A}}(b) = 1$. Now, E_{A} restricted to the Bauer simplex whose extreme set I consists of the sequence (a_n) and its limit

b_1, is an affine u.s.c. function, so it is E^{aff}, where E is a u.s.c. function on I. Since $E_{\text{A}} \le 1$, also $E \le 1$, and because

$$1 = E_{\text{A}}(b) = \sum_{n=1}^{\infty} 2^{-n} E(a_n),$$

E equals 1 at all points a_n ($n \ge 1$). By upper semi-continuity, $E(b_1) = 1$. But now $E_{\text{A}} - f_1$ is not u.s.c., because it has a "bad" jump at b_1. This concludes the example.

Example 4.8. Here is an example of $\mathcal{F}$ such that $\text{SUP}(\text{E}\mathcal{F}) > \sup\{\text{E}\mathcal{F}(\mu) \colon \mu \in \text{ex}K\}$. Consider a simplex whose extreme points consist of two elements b_1, b_2 and a sequence $(a_n)_{n\ge 2}$ converging to $b = \frac{1}{2}(b_1 + b_2)$. Set $f_1 = 1^{\text{aff}}_{\{b_1,b_2\}}$ (so that $f_1(b) = 1$), and $f_k = f_{k-1} + 1^{\text{aff}}_{\{a_k\}}$ ($k \ge 2$). Using Proposition 3.13, it is easy to verify $\text{SUP}(\text{E}\mathcal{F}) = 2$ (attained at b as in Example 2.19) and $\text{E}\mathcal{F} = 1$ at all extreme points.

Finally, we show the maximum of $\text{E}\mathcal{F}$ must be achieved on the *closure* of the extreme points, and we extend the estimate of Proposition 3.10 to the case of an affine sequence $\mathcal{F}$ defined on a Choquet simplex.

Theorem 4.9. *Suppose f is bounded. Then (1)* $\text{SUP}(\text{E}\mathcal{F}) = \sup\{\text{E}\mathcal{F}(\mu) \colon \mu \in \overline{\text{ex}K}\}$*, and (2) if $x \in \text{ex}K$ and $r(x)$ denotes its order within the set $\overline{\text{ex}K}$, then* $\text{E}\mathcal{F}(x) \le f(x) + r(x)u(x)$.

Remark 4.10. Unfortunately, the last inequality does not apply in general to $x \in \overline{\text{ex}K}$, and therefore does not provide an upper bound for $\text{SUP}(\text{E}\mathcal{F})$.

Proof. By Proposition 3.10, both assertions hold on Bauer simplices, where $(\text{E}\mathcal{F})|_{\text{ex}K}$ equals $\text{E}(\mathcal{F}|_{\text{ex}K})$. For a general Choquet simplex K let M denote the Bauer simplex of all probability measures supported by $\overline{\text{ex}K}$. The map π assigning to each such measure its barycenter is an affine continuous map from M onto K. Let $\mathcal{F}'$ denote the lift of $\mathcal{F}$ to M. If E is a superenvelope of $\mathcal{F}'$, then the pushed down function $E^{[K]}$ is a superenvelope of $\mathcal{F}$, and conversely a superenvelope of $\mathcal{F}$ lifts to a superenvelope of $\mathcal{F}'$. Therefore $\text{E}\mathcal{F} = (\text{E}\mathcal{F}')^{[K]}$ and

$$\begin{aligned}\text{SUP}(\text{E}\mathcal{F}) &\ge \text{SUP}(\text{E}\mathcal{F}|_{\overline{\text{ex}K}}) = \text{SUP}((\text{E}\mathcal{F}')^{[K]}|_{\overline{\text{ex}K}}) \\ &\ge \text{SUP}(\text{E}\mathcal{F}'|_{\text{ex}M}) = \text{SUP}(\text{E}\mathcal{F}') = \text{SUP}(\text{E}\mathcal{F}) \ .\end{aligned}$$

This proves (1). Now suppose $x \in \text{ex}K$, so x has a unique preimage x' in M, and $x' \in \text{ex}M$. Then

$$\text{E}\mathcal{F}(x) = (\text{E}\mathcal{F}')^{[K]}(x) = (\text{E}\mathcal{F}')(x') \le h(x) + r(x)u(x)$$

where the passage from x' to x in the last equality holds because the restriction of π to $\text{ex}M$ is a homeomorphism onto $\overline{\text{ex}K}$. □

5. The Sex Entropy Theorem

We now turn to symbolic extensions. For a self contained statement of our key result, we recall some notation.

Definitions 5.1. We let $K(X,T)$ denote the space of T-invariant Borel probabilities on X. We let $\varphi \colon (Y,S) \to (X,T)$ denote a symbolic extension (i.e. φ is a continuous surjection $Y \to X$ intertwining S and T, and (Y,S) is a subshift over a finite alphabet). The *extension entropy function* of φ is by definition the function $h^{\varphi}_{\text{ext}}\colon K(X,T) \to [0,+\infty)$ which sends μ to $\max\{h(S,\nu)\colon \nu \in K(Y,S) \text{ and } \varphi\nu = \mu\}$. Given (X,T), we will say a sequence of partitions P_k of X is *refining* if the maximum diameter of an element of P_k goes to zero with k and for each k the partition P_{k+1} refines P_k. The partitions have *small boundaries* if their boundaries have measure zero for all μ in $K(X,T)$.

Definition 5.2. An *entropy structure* for (X,T) is a sequence $\mathcal{H}$ of functions h_k defined on $K(X,T)$ in one of two ways.

(I) Suppose P_k is a refining sequence of finite Borel partitions with small boundaries. Define $\mathcal{H}$ by setting $h_k\colon \mu \mapsto h(T,\mu,P_k)$.

(II) Suppose (Z,R) is an aperiodic minimal zero entropy system, $\lambda \in K(Z,R)$, and $(X \times Z, T \times R)$ is the product system. Define $\mathcal{H}$ by first choosing an entropy structure $\mathcal{H}' = (h'_k)$ for $(X \times Z, T \times R)$ defined as in (I), and then setting $h_k(\mu) = h'_k(\mu \times \lambda)$. (The existence of such a structure $\mathcal{H}'$ follows from Fact 7.6(3).)

Remark 5.3. Not every system admits a refining sequence of finite Borel partitions with small boundaries, so it is tempting to take only (II) as a definition for entropy structure, with some fixed choice for (Z,R) such as the dyadic adding machine. However, this would not produce a canonical sequence $\mathcal{H}$ ($\mathcal{H}$ still depends on the choice of partition), and (I) is more direct in the many cases where it applies. The supply of sequences providing an entropy structure satisfying Theorem 5.5 will be greatly enriched in [D3].

Remark 5.4. In place of 5.2(II), it is natural to consider defining an entropy structure on $K(X,T)$ simply by pushing down to $K(X,T)$ (by supremum over preimages) the entropy structure $\mathcal{H}'$ of $(X \times Z, T \times R)$. This would produce affine u.s.c. functions on $K(X,T)$ which converge nondecreasingly to the entropy function. We did not make this choice because we also want the functions to have u.s.c. differences, and this need not be preserved by pushing down.

We now state our key result.

Theorem 5.5 (**Sex Entropy Theorem).** *Suppose T is a homeomorphism of a compact metrizable space X with entropy structure $\mathcal{H} = (h_k)$, and $\mathbf{h}_{\text{top}}(T) < \infty$. Suppose $E\colon K(X,T) \to [0,+\infty)$. Then the following are equivalent.*

1. *E is a bounded affine superenvelope of $\mathcal{H}$.*
2. *There is a symbolic extension $\varphi\colon (Y,S) \to (X,T)$ such that $E = h^{\varphi}_{\text{ext}}$.*

The proof of Theorem 5.5 splits into two parts. In the next section, we will prove the theorem for a convenient choice of $\mathcal{H}$ when the space X has dimension zero. Then in Section 7, we will address general (X,T) and alternate choices of $\mathcal{H}$.

6. Proof of the Sex Entropy Theorem: a zero dimensional case

This section is devoted to proving Theorem 5.5 in the following special case:

Case Z 6.1. X is zero dimensional and the entropy structure is given by 5.2(I) with each P_k a partition into nonempty clopen sets.

Notation 6.2. We now set some notation for the rest of this section. From the given clopen partitions P_k, we fix a presentation of (X,T) as the associated inverse limit of subshifts,

$$(X,T) = \overleftarrow{\lim_{k\to\infty}} (X_k, S) \ .$$

The alphabet for X_k is some set in bijective correspondence with P_k. The points of X_k correspond to itineraries through P_k under T. The bonding maps (factor maps) $X_{k+1} \to X_k$ are the amalgamations (surjective one-block maps) which reflect inclusion of an element of P_{k+1} in a unique element of P_k. The entropy structure $\mathcal{H}$ has an equivalent description here as $h_k(\mu) = h(S, \mu_k)$, where μ_k denotes the projection of μ to the system (X_k, S). Of course, $\lim_k h_k(\mu) = h(\mu)$, the entropy of T with respect to the measure μ. The functions h_k are nonnegative, u.s.c., affine and bounded by $h(T)$. As explained in [D2], and below, they also have u.s.c. differences.

Measure-theoretic conditional entropy.

Recall that a sequence (a_n) of real numbers is called *subadditive* if $a_{n+m} \le a_n + a_m$ for all $n, m \in \mathbb{N}$. For such a sequence (a_n), the sequence (a_n/n) converges to its infimum. Moreover, if (n_i) is a subsequence satisfying $n_i | n_{i+1}$ then the convergence is monotone along (n_i). By the way, we recall that the property $n_i | n_{i+1}$ is equivalent to (n_i) being a *base of an odometer*.

We recall some elementary entropy theory. For details and proofs see e.g. [Wa].

Let (Y, ν, S) be a measure preserving system, and let P_1 and P_2 be two partitions of Y, $P_1 \succcurlyeq P_2$ (P_1 a refinement of P_2). We denote

$$P_1^n := \bigvee_{i=0}^{n-1} T^{-i} P_1.$$

Throughout this paper we shall write:

$$H(\nu, P_1) = -\sum_{A \in P_1} \nu(A) \log(\nu(A));$$

$$H(\nu, P_1|P_2) = \sum_{B \in P_2} \nu(B) H(\nu_B, P_1) = H(\nu, P_1) - H(\nu, P_2),$$

where ν_B denotes the conditional measure induced by ν on B (or zero if $\nu(B) = 0$). It is known that $H(\nu, P_1^n | P_2^n)$ is a subadditive sequence, and hence divided by n converges to its infimum $h(\nu, P_1|P_2) = h(\nu, P_1) - h(\nu, P_2)$ known as the *conditional entropy.* We denote

$$H_n(\nu, P_1|P_2) = \frac{1}{n} H(\nu, P_1^n | P_2^n),$$

and we call it the (nth) *approximative conditional entropy.*

We can now recover a central observation of [D2]:

Proposition 6.3. *If φ is a factor map (quotient map) of subshifts, $\varphi\colon (Y, S) \to (Y', S)$, then the conditional entropy map $\nu \mapsto h(\nu) - h(\varphi\nu)$ is u.s.c.*

Proof. Let P_2 be the pullback of the zero coordinate partition of Y', and let P_1 be a generating clopen partition of Y which refines P_2. Then the function $H_n\colon \nu \mapsto H_n(\nu, P_1|P_2)$ is continuous. Therefore $\inf_n H_n$ is u.s.c. This infimum is the conditional entropy map. □

It follows that for our entropy structure $\mathcal{H}$, the difference functions $h_k - h_{k-1}$ are u.s.c. We can now prove the easy direction of the Sex Entropy Theorem.

Proof of Theorem 5.5 (2) $\implies$ (1) *in Case Z.*

For the given symbolic extension $\varphi\colon (Y, S) \to (X, T)$, we must show h_{ext}^{φ} is a bounded affine superenvelope of $\mathcal{H}$. Clearly h_{ext}^{φ} is bounded and each $h_{\text{ext}}^{\varphi} - h_k$ is nonnegative. For each k, postcomposing φ with the projection $X \to X_k$ produces a factor map φ_k of subshifts. Because

$$(h^{\varphi}_{\text{ext}} - h_k)(\mu) = \max\{h(\nu) - h(\varphi_k \nu) : \varphi\nu = \mu,\ \nu \in K(Y,S)\}$$

we have that $h^{\varphi}_{\text{ext}} - h_k$ is the pushdown (by φ) of a function which is u.s.c. (by Proposition 6.3) and affine. It follows that $h^{\varphi}_{\text{ext}} - h_k$ is u.s.c., and it is affine by Proposition 2.7 because the map φ sends ergodic measures to ergodic measures. □

Now we set some notation for the next lemma, which will be essential to our proof of the converse. Let (Y,S) and (Y',S) be full shifts: $Y = \Lambda^{\mathbb{Z}}$ and $Y' = \Lambda'^{\mathbb{Z}}$, where Λ and Λ' are finite sets (called *alphabets*). Suppose also that (Y',S) is a factor of (Y,S) via an amalgamation π. This means that the "zero coordinate partition" P_1 of Y into cylinders of the form $U_a = \{y \in Y : y(0) = a\}$ $(a \in \Lambda)$ is a refinement of the corresponding "zero coordinate partition" P_2 of Y' lifted to Y by preimage. In this case $H_n(\nu, P_1|P_2)$ is determined by observing the (cylinders generated by) blocks of length n in Y' and their preimages (which are blocks of length n in Y). Because P_1 and P_2 generate the Borel sigma-fields in Y and Y', respectively, they suffice for all evaluations of entropy, and we will not use any other partitions. We therefore skip their indicators in the notation of $h(\nu, P_1)$, $h(\nu, P_2)$, $h(\nu, P_1|P_2)$ and $H(\nu, P_1|P_2)$. Instead, we will write $h(\nu)$, $h(\pi\nu)$, $h(\nu|\pi\nu)$ and $H_n(\nu|\pi\nu)$, respectively, where $\pi\nu$ always denotes the image of ν on Y' defined by $\pi\nu(A) = \nu(\pi^{-1}(A))$. Now, suppose C is a block of some length m in Y. Its image by π is a block B in Y' of the same length. Consider the invariant measure ν_C carried by the periodic orbit of the sequence $\ldots CCC \ldots$. The code π sends this measure to the measure μ_B supported by $\ldots BBB \ldots$. In such a case, we define

$$H_n(C|B) = H_n(\nu_C|\mu_B).$$

Lemma 6.4. *Let (Y,S) and (Y',S) be full shifts, $Y' = \Lambda'^{\mathbb{Z}}$ and $Y = \Lambda^{\mathbb{Z}}$. Let $\pi\colon Y \to Y'$ be an amalgamation. Denote $\lambda = \#\Lambda$. For $n, m \in \mathbb{N}$ and blocks $C \in \Lambda^m$ and $B \in \Lambda'^m$, $B = \pi(C)$, let $H_n(C|B)$ denote the corresponding approximative conditional entropy. Then for every $n \in \mathbb{N}$ and $\epsilon > 0$ there exists an $m_{(n,\epsilon)} \in \mathbb{N}$ such that for every $m \geq m_{(n,\epsilon)}$ the following holds*

$$\sum_{C \in \pi^{-1}(B)} \exp[-mH_n(C|B)] \leq me^{m\epsilon}(\log\lambda).$$

Proof. Consider $\pi^{-1}(B)$ as a probability space with uniform probability $P(C) = \frac{1}{\#\pi^{-1}(B)} = p \geq \lambda^{-m}$, with a random variable $X(C) = H_n(C|B)$. The expression to estimate equals $\frac{1}{p}\mathrm{E}(e^{-mX})$ (here E denotes the expected value). By Lemma 2 in [D2] (which is simply a conditional version of Lemma 1 in [BGH]), we have for every $t \geq 0$ that

$$\#\{C \in \pi^{-1}(B) \colon H_n(C|B) \leq t\} \leq e^{m(t+\epsilon)}$$

for m sufficiently large. Clearly, the above cardinality does not exceed $\#\pi^{-1}(B)$, and it is zero for $t < 0$. Eventually, we can write

$$p\#\{C \in \pi^{-1}(B) \colon H_n(C|B) \leq t\} \ \leq\ \begin{cases} 0\ , & t < 0; \\ \min(pe^{m(t+\epsilon)}, 1)\ , & t \geq 0. \end{cases}$$

This can be interpreted as saying that the cumulative distribution function of X, $F_X(t) = \text{Prob}(X \le t)$, is below the cumulative distribution function of W, where W is a random variable with density

$$f(t) = pme^{m(t+\epsilon)}$$

on the interval $[0, \frac{-\log p}{m} - \epsilon]$, and with an atom of mass $pe^{m\epsilon}$ at zero. This means that the mass of probability of the distribution of X is moved to the right compared to that of W. Because the function e^{-mx} is decreasing, the expected value $\mathrm{E}(e^{-mX})$ is not larger than

$$\mathrm{E}(e^{-mW}) = pme^{m\epsilon}(\tfrac{-\log p}{m} - \epsilon) + 1 \cdot pe^{m\epsilon} \le pme^{m\epsilon}(\log \lambda + \frac{1}{m} - \epsilon).$$

This provides the required estimate of $\frac{1}{p}\mathrm{E}(e^{-mX})$. □

Topological fiber entropy.

For the final entropy calculation of our theorem (Stage 4), we will need a variational principle which lets us make the transition from word counts to the sex entropy of a measure. To recall this principle, let (X, T) and (Y, S) be two topological dynamical systems and let $\pi \colon Y \to X$ be a topological factor map. For an open cover $\mathcal{A}$ of Y and an x in X we set

$$\mathbf{H}(\mathcal{A}|x) = \log \min\{\#\mathcal{F} \colon \mathcal{F} \subset \mathcal{A}, \bigcup \mathcal{F} \supset \pi^{-1}(x)\}$$

and for a probability measure μ on X we set

$$\mathbf{H}(\mathcal{A}|\mu) = \int \mathbf{H}(\mathcal{A}|x) d\mu(x) \ .$$

The function $x \mapsto \mathbf{H}(\mathcal{A}|x)$ is easily seen to be u.s.c. on X and hence also $\mu \mapsto \mathbf{H}(\mathcal{A}|\mu)$ is a u.s.c. function on the space of measures, and it is obviously affine. Moreover, if the cover $\mathcal{A}$ is a partition into closed and open sets, then these maps are continuous. If μ is an invariant measure then, by a subadditivity argument, the sequence $\frac{1}{n}\mathbf{H}(\mathcal{A}^n|\mu)$ (where $\mathcal{A}^n$ denotes the joint cover $\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{A})$) converges to its infimum and the convergence is nonincreasing along any subsequence $(p_k)_{k\in\mathbb{N}}$ such that for each k, p_{k+1} is a multiple of p_k. Thus, on $K(X,T)$ we have a well defined affine u.s.c. function

$$\mathbf{h}(\mathcal{A}|\cdot) = \lim_n \frac{1}{n}\mathbf{H}(\mathcal{A}^n|\cdot)$$

which we call the topological fiber entropy function on K with respect to $\mathcal{A}$. Finally, on K we define the function

$$\mathbf{h}(Y|\cdot) = \sup_{\mathcal{A}} \mathbf{h}(\mathcal{A}|\cdot)$$

which we call *the topological fiber entropy function.* We will need the following fact: for $\mu \in K(X,T)$,

$$\mathbf{h}(Y|\mu) = \sup_{\nu}\{h(\nu|\mu) \colon \nu \in K(Y,S) \text{ and } \pi\nu = \mu\} \ . \tag{6.5}$$

In other words, $\mathbf{h}(Y|\mu) = h^{\pi}_{\text{ext}}(\mu) - h(\mu)$. This fact is essentially a special case of the "Relativised Variational Principle" [LeWa, Theorem 2.1]. It is also the "Inner Variational Principle", which holds for not necessarily metrizable systems [DS1, Theorem 4]. In the special case that (Y, S) is a subshift, it is known that the supremum in the definition of $\mathbf{h}(Y|\cdot)$ is attained for the zero coordinate partition

$\mathcal{A}$ (which is also a cover). We summarize what we will need for our entropy computation in the following lemma.

Lemma 6.6. *Suppose $\varphi\colon (Y,S)\to(X,T)$ is a factor map, (Y,S) is a subshift, and $\mathcal{A}$ is the zero coordinate partition of Y. Then for $\mu\in K(X,T)$,*

$$h^{\varphi}_{\text{ext}}(\mu)=\lim_n \frac{1}{n}\mathbf{H}(\mathcal{A}^n|\mu)+h(\mu)$$

and the functions $\mu\mapsto \mathbf{H}(\mathcal{A}^n|\mu)$ are continuous on $K(X,T)$.

We are now ready for the difficult direction of the Sex Entropy Theorem.

Proof of Theorem 5.5 (1) $\implies$ (2) *in Case Z.*

Given the bounded affine superenvelope E, we will construct a symbolic extension φ such that $h^{\varphi}_{\text{ext}}=E$. The construction is a combination and simplification of methods from [BFF] and [D2] which at the same time allows for better control over entropy properties of the extension. This new construction splits into two main stages. From an affine superenvelope we derive an SWO, a simplified version of the notion of a word oracle introduced in [BFF]; then, from the SWO we construct a symbolic extension with the desired sex entropy. In fact, we present these steps in the reversed order as stages 2 and 3.

STAGE 1. *Preliminary reshaping*

As in [D2], we begin with the observation, that it is convenient to consider the direct product of (X,T) with an odometer. Let $\mathbf{p}=(p_k)_{k\in\mathbb{N}}$ be a sequence of positive integers such that $p_{k+1}=p_kq_k$ with q_k being an integer larger than 1, for each k. Let $(G_{\mathbf{p}},\tau)$ denote the rotation by 1 of the odometer to base $\mathbf{p}$. We denote $(X',T')=(X\times G_{\mathbf{p}},T\times\tau)$.

We represent $(X,T)=\overleftarrow{\lim}(X_k,S)$ as the shift transformation on a closed shift-invariant family of infinite matrices

$$x=(x_{k,n})_{k\in\mathbb{N},n\in\mathbb{Z}},$$

such that each row (indexed by k) is a sequence over some finite alphabet Δ_k. We let

$$x_k=(x_{s,n})_{s\le k,n\in\mathbb{Z}}$$

denote the projection of x onto the first k rows. In this setting, (X_k,S) is the symbolic system obtained from (X,T) by the projection onto the first k rows. The alphabet of X_k is contained in the product $\Lambda_k=\Delta_1\times\cdots\times\Delta_k$. Note that each projection from X_{k+1} to X_k is an amalgamation.

The product (X',T') is realized in such manner that in each row x_k we apply a p_k-periodically repeating marker, say a comma, and the positions of the commas in the $(k{+}1)$st row coincide with the positions of some (every q_kth) commas in the kth row. (Formally, the alphabet Δ'_k of the kth row is $\Delta_k\cup\Delta_k\times\{,\}$). Now, the projection from X'_k to X_k (erasing the commas in the first k rows) induces a map from $K'_k=K(X'_k,S)$ onto $K_k=K(X_k,S)$ and even though this map is not 1-1, it is obvious that $h(\mu'_k)=h(\mu_k)$ whenever μ'_k projects to μ_k. Thus the conditional entropy functions on $K'=K(X',T')$ coincide with the functions h_k lifted from $K=K(X,T)$. So, if E_{A} is a bounded affine superenvelope for $\mathcal{H}$ on K, its lift (also denoted by E_{A}) becomes a bounded affine superenvelope for the entropy sequence (also denoted by $\mathcal{H}$) on K'. So, it suffices to construct a symbolic extension of (X',T') with the required sex entropy function. From now on we will use (X,T)

to denote (X',T'), we will also skip the apostrophes in X'_k, K'_k, etc. We are free to choose the base $\mathbf{p}$ according to our needs, and we will specify $\mathbf{p}$ in Stage 3.

For given k, the blocks B of length p_k over the alphabet Δ_k ending with a comma, will be called k-blocks and their collection will be denoted by $\mathcal{B}_k$. A rectangular $p_k \times k$ block R having a concatenation of $\frac{p_k}{p_s}$ of s-blocks in the sth row (for each $s = 1, 2, \dots k$) will be called an k-rectangle, and the collection of such will be denoted by $\mathcal{R}_k$. Notice that a $(k{+}1)$-rectangle is a concatenation of q_k k-rectangles with a $(k{+}1)$-block added as a new row, which we denote by

$$R = \begin{pmatrix} R^{(1)} R^{(2)} \dots R^{(q_k)} \\ B \end{pmatrix}$$

(see figure below; stars replace any symbols from any Δ_k).

```
***, ***, ***, ***, ***, ***, ***, ***, ***, ***, ***, ***,
***  ***  ***  ***, ***  ***  ***  ***, ***  ***  ***  ***,
***  ***  ***  ***  ***  ***  ***  ***  ***  ***  ***  ***,
```

Also note that a k-rectangle is in fact a block over Λ_k (with added commas). Let $\mathcal{R}_k(X)$ denote the collection of all k-rectangles that actually occur in X.

STAGE 2. *Simplified word oracle.*

By a *simplified word oracle (SWO)* we shall understand a sequence of functions $\alpha_k \colon \mathcal{R}_k \mapsto \mathbb{N}$ such that $\alpha_k(R) = 0$ if and only if $R \notin \mathcal{R}_k(X)$, and

$$\alpha_k(R^{(1)})\alpha_k(R^{(2)}) \cdots \alpha_k(R^{(q_k)}) \ge \sum_{B \in \mathcal{B}_{k+1}} \alpha_{k+1} \begin{pmatrix} R^{(1)} R^{(2)} \dots R^{(q_k)} \\ B \end{pmatrix}. \tag{6.7}$$

We now describe a construction of a symbolic extension (Y,S) of (X,T) given an SWO. Let $\overline{X}_k$ denote the set of all possible infinite concatenations of k-rectangles from $\mathcal{R}_k(X)$. Clearly, $X_k \subset \overline{X}_k$. Our starting point will be constructing (in a consistent way), for each k, a certain symbolic extension $(\overline{Y}_k, S)$ of $(\overline{X}_k, S)$ such that precisely $\alpha_k(R)$ blocks will be seen above each k-rectangle R.

Note that $\mathcal{R}_1 = \mathcal{B}_1$. Let $\lambda \in \mathbb{N}$ be such that $\lambda^{p_1} \ge \sum_{R \in \mathcal{R}_1} \alpha_1(R)$, and let Λ_Y be an alphabet with cardinality λ.

Step 1. For each $R \in \mathcal{R}_1$ pick a family $\mathcal{F}_1(R)$ of $\alpha_1(R)$ different blocks of length p_1 over Λ_Y (for R not appearing in X this family is empty). The cardinality of Λ_Y allows to do it so that for different blocks R the families $\mathcal{F}_1(R)$ are disjoint. Let $\mathcal{F}_1$ be the union of all families $\mathcal{F}_1(R)$. Attach a comma at the end of every block in $\mathcal{F}_1$. By disjointness, we can define a map ρ_1 from $\mathcal{F}_1$ onto $\mathcal{R}_1(X)$ sending each C to the unique R for which $C \in \mathcal{F}_1(R)$. Let $\overline{Y}_1$ be the set of all possible concatenations of blocks from $\mathcal{F}_1$. Clearly, ρ_1 sends $\overline{Y}_1$ onto $\overline{X}_1$ and what can be seen above each 1-rectangle R is $\mathcal{F}_1(R)$ whose cardinality is exactly $\alpha_1(R)$.

Step $k{+}1$. Suppose the task has been completed for some k. Consider a concatenation $R^{(1)} R^{(2)} \dots R^{(q_k)}$ of some k-rectangles. Above it in $\overline{Y}_k$, exactly as many as $\alpha_k(R^{(1)})\alpha_k(R^{(2)}) \cdots \alpha_k(R^{(q_k)})$ blocks can be seen: namely all possible concatenations of q_k blocks: first from $\mathcal{F}_k(R^{(1)})$, next from $\mathcal{F}_k(R^{(2)})$, and so on. By the inequality (6.7) in the definition of an SWO, it is thus possible, for each $(k{+}1)$-rectangle $R = \begin{pmatrix} R^{(1)} R^{(2)} \dots R^{(q_k)} \\ B \end{pmatrix}$, to disjointly select from these concatenations a family $\mathcal{F}_{k+1}(R)$ of cardinality $\alpha_{k+1}(R)$. Note that if two $(k{+}1)$-rectangles differ already in the the first k rows then their families are chosen from disjoint collections of concatenations, so obviously they are disjoint. Let $\mathcal{F}_{k+1}$ be the union of all so selected families (over all $(k+1)$-rectangles R, after removing all but the terminal comma in every such block). Let ρ_{k+1} be the map from $\mathcal{F}_{k+1}$ onto $\mathcal{R}_{k+1}(X)$ sending each C to the unique R for which $C \in \mathcal{F}_{k+1}(R)$. Let $\overline{Y}_k$ be the set of all possible concatenations of blocks from $\mathcal{F}_{k+1}$. Clearly, ρ_{k+1} sends $\overline{Y}_{k+1}$ onto $\overline{X}_{k+1}$ and what can be seen above each $(k+1)$-rectangle R is $\mathcal{F}_{k+1}(R)$ whose cardinality is exactly $\alpha_{k+1}(R)$. The task has been now completed for $k+1$.

Notice that there are natural maps $\pi_k \colon \overline{X}_{k+1} \to \overline{X}_k$ (by projection onto the first k rows, hence they coincide on X_{k+1} with the bonding maps of the inverse limit defining X). These maps are not surjective, but it will not be essential. Also, we have natural (not surjective) maps $\kappa_k \colon \overline{Y}_{k+1} \to \overline{Y}_k$. Namely, each $y \in \overline{Y}_{k+1}$ being a concatenation of blocks from $\mathcal{F}_{k+1}$ is, at the same time, a concatenation of smaller blocks from $\mathcal{F}_k$, hence, adding appropriate commas, we can view y as an element of $\overline{Y}_k$. Now notice that the diagram below commutes:

$$\begin{array}{ccccccc}
\overline{Y}_1 & \underset{\kappa_1}{\longleftarrow} & \overline{Y}_2 & \underset{\kappa_2}{\longleftarrow} & \overline{Y}_3 & \underset{\kappa_3}{\longleftarrow} & \dots \\
\Big\downarrow{\scriptstyle \rho_1} & & \Big\downarrow{\scriptstyle \rho_2} & & \Big\downarrow{\scriptstyle \rho_3} & & \\
\overline{X}_1 & \underset{\pi_1}{\longleftarrow} & \overline{X}_2 & \underset{\pi_2}{\longleftarrow} & \overline{X}_3 & \underset{\pi_3}{\longleftarrow} & \dots
\end{array}$$

The above diagram defines a factor map ρ from the inverse limit $(\overline{Y}, S) = \overleftarrow{\lim}(\overline{Y}_k, S)$ onto the inverse limit $(\overline{X}, S) = \overleftarrow{\lim}(\overline{X}_k, S)$. Clearly $X \subset \overline{X}$ and we define Y to be the preimage by ρ of X. Because all maps κ_k do not alter the symbols of y_{k+1} (only more commas are added), $(\overline{Y}, S)$ (and hence also (Y, S)) can be viewed as a subsystem of the direct product of a symbolic system (over the alphabet Λ_Y) with the odometer to base p. Finally, the odometer can be replaced by a zero entropy strictly ergodic symbolic extension, for example, by a regular Toeplitz flow of which the odometer is a factor [JK]. This completes the construction of a symbolic extension of (X, T) using a given SWO.

We have the following convergence:

Lemma 6.8. *For $\mu \in K$ and μ_k denoting the projection of μ to K_k,*

$$\mathbf{h}(Y|\mu) = \lim_k \mathbf{h}(\overline{Y}_k|\mu_k). \tag{6.9}$$

Proof. Given $k \in \mathbb{N}$, we will justify the following claims:

$$\begin{aligned}
\mathbf{h}(Y|\mu) &= \sup\{h(\nu) - h(\mu) : \nu \in \rho^{-1}(\mu)\} \\
&\le \sup\{h(\kappa_k(\nu)) - h(\mu_k) : \nu \in \rho^{-1}(\mu)\} \\
&\le \sup\{h(\nu_k) - h(\mu_k) : \nu_k \in \rho_k^{-1}(\mu_k)\} \ = \ \mathbf{h}(\overline{Y}_k|\mu_k) \\
&= \max\{h(\nu_k) - h(\mu_k) : \nu_k \in \rho_k^{-1}(\mu_k)\} \ .
\end{aligned}$$

The first equality follows from the variational fact (6.5), because $\rho(\nu) = \mu$ implies ν is supported on Y. The first inequality holds because for all k we have $h(\mu_k) \le h(\mu)$ and also $h(\kappa_k(\nu)) = h(\nu)$ (the maps κ_k are "essentially" embeddings, only adding periodic commas). For the second inequality, we take a supremum over a larger set. The second equality is again (6.5). The supremum is achieved by Proposition 6.3.

Similarly, $\mathbf{h}(\overline{Y}_k|\mu_k)$ is nonincreasing in k. We conclude $\mathbf{h}(Y|\mu) \le \lim_k \mathbf{h}(\overline{Y}_k|\mu_k)$.

For the converse inequality, let M_k denote the set of measures ν_k in $\rho_k^{-1}(\mu_k)$ such that $h(\nu_k) - h(\mu_k) \ge \lim_\ell \mathbf{h}(\overline{Y}_\ell|\mu_\ell)$. These sets are nonempty by the mentioned monotonicity, since some measure ν_k satisfies $h(\nu_k)-h(\mu_k)=\mathbf{h}(\overline{Y}_k|\mu_k)$. These sets are also closed, by upper semi-continuity of h in symbolic systems, and $\kappa_k(M_{k+1})\subset M_k$. By compactness, we can thus find a measure ν on $\prod_k \overline{Y}_k$ such that $\kappa_k(\nu)\in M_k$ for each k. This measure ν is supported on Y, so $\mathbf{h}(Y|\mu)\geq h(\nu)-h(\mu)\geq \lim_k \mathbf{h}(\overline{Y}_k|\mu_k)$. □

STAGE 3. *Constructing an SWO from an affine superenvelope.*

Let E_{A} be an affine superenvelope of $\mathcal{H}$. By definition the functions $E_k := E_{\mathrm{A}} - h_k$ are nonnegative affine and u.s.c. They converge nonincreasingly to the function $E_{\mathrm{A}} - h$, which therefore is also nonnegative, affine and u.s.c.

Before we continue with the proof, we need one more technical lemma:

Lemma 6.10. *Define functions $E_k^{[K_k]}$ on $K(X,T)$ as follows:*

$$E_k^{[K_k]}(\mu) = \sup_{\xi\in K:\ \xi_k=\mu_k} E_k(\xi)\ .$$

Then these are affine u.s.c. functions and $\lim_k E_k^{[K_k]} = \lim_k E_k = E_{\mathrm{A}} - h$.

Proof. Given k, the function $E_k^{[K_k]}$ is defined by supremum over preimages under the natural projection $K(X,T)\to K(X_k,S)$. This projection is a continuous affine surjection between Choquet simplices, and it sends extreme points to extreme points because it sends ergodic measures to ergodic measures. Because E_k is affine, the affinity of $E_k^{[K_k]}$ then follows from Proposition 2.7.

Clearly the sequence $E_k^{[K_k]}$ converges pointwise, as a nonincreasing sequence of nonnegative functions, and $\lim_k E_k^{[K_k]} \geq \lim_k E_k$. So, only one inequality for the limits requires verification. Suppose

$$\lim_k E_k^{[K_k]}(\mu) > \lim_k E_k(\mu) + \epsilon.$$

For each k we can pick a measure ξ^k in K projecting to μ_k such that $E_k^{[K_k]}(\mu) = E_k(\xi^k)$. It is seen that the ξ^k approach μ as $k\to\infty$, because they agree with μ on finer and finer clopen partitions. Further, for $s<k$, we have $E_s(\xi^k)\geq E_k(\xi^k)$. Leaving s fixed and letting $k\to\infty$, we obtain, by upper semi-continuity, $E_s(\mu)\geq \lim_k E_k(\xi^k) > \lim_k E_k(\mu)+\epsilon$, holding for each s, which is impossible. □

We can now continue with the main proof. By Definition 2.1(2) and Theorem 2.11, we can represent each function $E_k^{[k]}$ as a strictly decreasing limit of affine continuous functions $f_{k,i}$ defined on K_k and then lifted to K. Moreover, since $E_k^{[k]}$ is nonincreasing in k, we can easily arrange the double sequence $f_{k,i}$ to be nonincreasing also in k for each i. Then the sequence of functions $f_k := f_{k,k}$ satisfies:

(a) $f_k > E_k^{[K_k]}$, for each $k\in\mathbb{N}$, and
(b) $\lim_k f_k = E_{\mathrm{A}} - h$.

Because $f_k - (h_{k+1}-h_k) > E_k - (h_{k+1}-h_k) = E_{k+1}$ and the function on the left is constant on the fibers of the projection $K\to K_{k+1}$, we have

$$f_k - (h_{k+1}-h_k) > E_{k+1}^{[K_{k+1}]}.$$

Using the Sandwich Theorem 2.8 (on K_{k+1}), we can find a continuous function f'_{k+1} with $f_k - (h_{k+1}-h_k) > f'_{k+1} > E_{k+1}^{[K_{k+1}]}$. Then, $\min(f_{k+1}, f'_{k+1})$ is a continuous function strictly above $E_{k+1}^{[K_{k+1}]}$, hence, by Theorem 2.11, we can find an affine continuous function f''_{k+1} such that $f_k - (h_{k+1} - h_k) > f''_{k+1} > E_{k+1}^{[K_{k+1}]}$. We redefine f_{k+1} to equal f''_{k+1}. We do this for $k = 1$ and we proceed inductively.

In the end, the conditions (a) and (b) still hold, and $f_k - (h_{k+1} - h_k) > f_{k+1}$. Appealing to upper semicontinuity of $h_{k+1} - h_k$, we choose a sequence (ϵ_k) of positive numbers decreasing to zero and satisfying

(c) $f_k - f_{k+1} - 5\epsilon_k > h_{k+1} - h_k$.

Next, we will choose the base $\mathbf{p} = (p_k)$ of an odometer, i.e., a sequence satisfying $p_k | p_{k+1}$, growing rapidly enough to satisfy several conditions. Namely, note that if (n_i) is such a base then, for each k, the functions $H_{n_i}(\mu_{k+1}|\mu_k)$ converge nonincreasingly to $h_{k+1} - h_k$, and thus by (c) the continuous functions $\mu \mapsto \max\{f_k - f_{k+1} - 5\epsilon_k,\ H_{n_i}(\mu_{k+1}|\mu_k)\}$ converge nonincreasingly to the continuous function $f_k - f_{k+1} - 5\epsilon_k$. By Property 2.2(4) this convergence is uniform, so we may choose p_k to be an index n_i large enough to guarantee

(d) $(f_k - f_{k+1})(\mu) - 4\epsilon_k > H_{p_k}(\mu_{k+1}|\mu_k)$ for all μ

and we also require

(e) $p_{k+1} \geq m_{(p_k,\epsilon_k)}$ of Lemma 6.4 applied to the amalgamation from X_{k+1} to X_k,

(f) $p_{k+1} e^{-p_{k+1}\epsilon_k} \log(\#\Lambda_{k+1}) \leq 1$, and

(g) $e^{p_{k+1}(b+\epsilon_k)} \geq \lceil e^{p_{k+1}b} \rceil$ whenever $b \geq 0$, where $\lceil\cdot\rceil$ denotes the integer ceiling. (It suffices to require $e^{p_{k+1}\epsilon_k} \geq 2$.)

Observe that K_{k+1} is a subset of the set of all shift invariant measures of the full shift over the alphabet Λ_{k+1}. The functions $h_{k+1} - h_k$ and $\mu \mapsto H_{p_{k+1}}(\mu_{k+1}|\mu_k)$ are well defined (as entropy) and u.s.c. on this larger set. For each k we find a continuous affine extension of f_k to measures on the full shift over the alphabet Λ_k, and by lifting we also have f_k defined on measures on the full shift over the alphabet Λ_{k+1}. Then (d) holds on some open neighborhood U_{k+1} of K_{k+1}.

Also note that the periodic measure μ_B carried by the orbit of $\dots BBB \dots$, where B is any sufficiently long block appearing in X_k, is in U_k. As explained prior to Lemma 6.4, we will identify the block with the measure, so that $h_k(B)$, $f_k(B)$, etc. have a definite meaning. Because a subsequence of a base of an odometer is again a base of an odometer, and the inequalities (d) – (g) are satisfied for subsequences of (p_k), we may specify two more requirements on the speed of growth of the numbers p_k. First, take the length p_k of a k-rectangle so large that

(h) if R is a k-rectangle, then μ_R is in U_k.

Then, by appeal to f_k being continuous affine, we take p_k large enough that the value of f_k on any concatenation $R^{(1)}R^{(2)}\dots R^{(q)}$ of k-rectangles is close to the corresponding convex combination of values:

(i) $$\left| f_k(R^{(1)}R^{(2)}\dots R^{(q)}) - \tfrac{1}{q}\sum_{i=1}^{q} f_k(R^{(i)}) \right| < \epsilon_k \ .$$

We can now define the SWO. For a k-rectangle R appearing in X, let

$$\alpha_k(R) = \lceil e^{p_k f_k(R)} \rceil \ .$$

We need to verify the condition (6.7) in the definition of an SWO. Let $C = R^{(1)}R^{(2)}\dots R^{(q_k)}$ be a concatenation of k-rectangles appearing in X. If C does not occur in X, then the right side of (6.7) is zero and the inequality (6.7) holds trivially; so suppose C occurs in X. Then, using (g), (d), and (i) for the consecutive inequalities, we can write

$$\sum_{B\in\mathcal{B}_{k+1}} \alpha_{k+1}\binom{C}{B} = \sum_{B\in\mathcal{B}_{k+1}} \left\lceil \exp\left[p_{k+1} f_{k+1}\binom{C}{B}\right]\right\rceil$$
$$\leq \sum_{B\in\mathcal{B}_{k+1}} \exp\left[p_{k+1}\left(f_{k+1}\binom{C}{B} + \epsilon_k\right)\right]$$
$$\leq \sum_{B\in\mathcal{B}_{k+1}} \exp\left[p_{k+1}\left(f_k(C) - H_{p_k}\left(\binom{C}{B}\Big|C\right) - 3\epsilon_k\right)\right]$$
$$\leq \exp\left[\tfrac{p_{k+1}}{q_k}\sum_{i=1}^{q_k} f_k(R^{(i)})\right] e^{-2p_{k+1}\epsilon_k} \sum_{B\in\mathcal{B}_{k+1}} \exp\left[-p_{k+1}H_{p_k}\left(\binom{C}{B}\Big|C\right)\right].$$

The first term in the last expression above is not larger than

$$\alpha_k(R^{(1)})\alpha_k(R^{(2)})\cdots\alpha_k(R^{(q_k)}).$$

By (e) and Lemma 6.4, the remaining part of that expression is not larger than $p_{k+1}e^{-p_{k+1}\epsilon_k}\log(\#\Lambda_{k+1})$, which by (f) is at most 1. This verifies (6.7).

STAGE 4. *Entropy calculation*

It remains to verify $h^{\varphi}_{\text{ext}} = E_A$. First we compute $\mathbf{h}(\overline{Y}_k|\mu_k)$ for measures μ_k carried by $\overline{X}_k$. To begin, consider a periodic measure μ_C, where $C = R^{(1)}R^{(2)}\dots R^{(ql)}$ is a concatenation of ql k-rectangles ($q, l \in \mathbb{N}$). Let $\mathcal{B}(i,n,C)$ denote the set of blocks of length n occurring as the word $y[i, i+n-1]$ for some point y such that $(\rho_k(y))[jqlp_k, (j+1)qlp_k - 1] = C$ for all j (i.e., $\rho_k(y) = \dots CCC\dots$ with $y[0, qlp_k - 1]$ sitting directly above C). Denoting by $\mathcal{A}_k$ the zero coordinate partition of $\overline{Y}_k$, we have

$$\mathbf{H}(\mathcal{A}_k^{lp_k}|\mu_C) = \frac{1}{qlp_k}\sum_{i=1}^{qlp_k}\log\#\mathcal{B}(i, lp_k, C)\ .$$

The above approximately equals (the error estimate is given with the $\pm$ sign)

$$\frac{1}{qlp_k}\sum_{i=1}^{ql} p_k[\log\#\mathcal{B}(ip_k, lp_k, C) \pm\ p_k\log\lambda]$$
$$= \left[\frac{1}{ql}\sum_{i=1}^{ql}\sum_{j=0}^{l-1}\log\alpha_k(R^{(i+j)})\right]\ \pm\ p_k\log\lambda\ ,\quad \text{with } i+j \text{ regarded modulo } q_l,$$
$$= \left[\frac{1}{ql}\sum_{i=1}^{ql} l\log\alpha_k(R^{(i)})\right]\ \pm\ p_k\log\lambda$$
$$= \left[l\frac{1}{ql}\sum_{i=1}^{ql}\log\left(\lceil\exp(p_k f_k(R^{(i)}))\rceil\right)\right]\ \pm\ p_k\log\lambda$$
$$= \left[l\frac{1}{ql}\sum_{i=1}^{ql}\log\left(\exp(p_k[f_k(R^{(i)}) \pm \epsilon_{k-1}]\right)\right]\ \pm\ p_k\log\lambda\ ,\quad \text{using (g)},$$
$$= \left[lp_k\frac{1}{ql}\sum_{i=1}^{ql} f_k(R^{(i)})\right]\ \pm\ (p_k\log\lambda + lp_k\epsilon_{k-1})\ .$$

Now, by (i), we can replace the above by

$$lp_k f_k(\mu_C) \,\pm\, \Big(p_k \log\lambda + lp_k\,(\epsilon_{k-1}+\epsilon_k)\Big).$$

Thus

$$\frac{1}{lp_k}\mathbf{H}(\mathcal{A}_k^{lp_k}|\mu_C) = f_k(\mu_C) \,\pm\, \left(\frac{\log\lambda}{l} + \epsilon_{k-1}+\epsilon_k\right)\ .$$

For fixed l, measures of the form μ_C with $C = R^{(1)}R^{(2)}\dots R^{(ql)}$ are dense. It follows from continuity of $\mathbf{H}(\mathcal{A}_k^{lp_k}|\cdot)$ (Lemma 6.6) that the last formula holds for all invariant measures μ_k on $\overline{X}_k$. Passing with l to infinity we obtain

$$\mathbf{h}(\overline{Y}_k|\mu_k) = f_k(\mu_k) \,\pm\, (\epsilon_{k-1}+\epsilon_k)\ .$$

We now pass to the factor map π between Y and X. Consider an invariant measure μ on X. By projection to $\overline{X}_k$ this measure produces a sequence of measures μ_k supported by X_k. By Lemma 6.8, since $f_k(\mu) = f_k(\mu_k)$, and by (b), we arrive at

$$h^{\varphi}_{\text{ext}}(\mu) = h(\mu) + \mathbf{h}(Y|\mu) = h(\mu) + \lim_k f_k(\mu) = E_{\text{A}}(\mu)$$

which concludes the proof of the Sex Entropy Theorem in Case Z. □

7. Proof of the Sex Entropy Theorem: the general case

To complete the proof of the Sex Entropy Theorem, we will lift general systems to zero dimensional systems by *principal extensions.*

Definition 7.1. [Le] An extension $\varphi\colon\ (X',T') \to (X,T)$ is a *principal extension* if $h(\nu) = h(\varphi\nu)$ for every $\nu \in K' = K(X',T')$.

Lemma 7.2. *Let $\varphi\colon (Y,S) \to (X,T)$ be an extension of (X,T) and let $\varphi'\colon (X',T') \to (X,T)$ be a principal extension of (X,T). Then the fiber product (Z,S') of these extensions, defined as*

$$Z = \{(y,x')\colon \varphi(y) = \varphi'(x')\}$$

with $S' = (S\times T')|_Z$, is a principal extension of (Y,S).

Proof. Z projects naturally to both Y and X' so that the following diagram commutes:

$$\begin{array}{ccccc} & & Z & & \\ & \swarrow & & \searrow & \\ Y & & & & X' \\ & \searrow & & \swarrow & \\ & & X & & \end{array}$$

In this proof it will be convenient to use a classical notation for measure-theoretic entropies and conditional entropies with respect to given σ-fields [Pa]. Let ν denote an invariant measure supported by Z, and let $\mathcal{B}_Z$, $\mathcal{B}_Y$, $\mathcal{B}_{X'}$ and $\mathcal{B}_X$ denote the Borel σ-fields of the respective spaces, lifted to Z. Then, because $B_{X'} \succcurlyeq \mathcal{B}_X$, we have

$$h_\nu(\mathcal{B}_Y|\mathcal{B}_{X'}) \le h_\nu(\mathcal{B}_Y|\mathcal{B}_X).$$

On the other hand, since X' is a principal extension of X,

$$h_\nu(\mathcal{B}_Y|\mathcal{B}_{X'}) = h_\nu(\mathcal{B}_Y \vee \mathcal{B}_{X'}|\mathcal{B}_X) - h_\nu(\mathcal{B}_{X'}|\mathcal{B}_X) \ge h_\nu(\mathcal{B}_Y|\mathcal{B}_X) - 0\ .$$

So, $h_\nu(\mathcal{B}_Y|\mathcal{B}_{X'}) = h_\nu(\mathcal{B}_Y|\mathcal{B}_X)$, and hence

$$\begin{aligned} h_\nu(\mathcal{B}_Z) &= h_\nu(\mathcal{B}_Y \vee \mathcal{B}_{X'}) = h_\nu(\mathcal{B}_Y|\mathcal{B}_{X'}) + h_\nu(\mathcal{B}_{X'}) \\ &= h_\nu(\mathcal{B}_Y|\mathcal{B}_X) + h_\nu(\mathcal{B}_X) = h_\nu(\mathcal{B}_Y \vee \mathcal{B}_X) = h_\nu(\mathcal{B}_Y)\ . \end{aligned}$$

□

Remarks 7.3. Asymptotic h-expansiveness was introduced by Misiurewicz [M1] as a sufficient condition for the existence of a measure of maximal entropy. For background, see [M1, M2, BFF, DGS] and their references. For simplicity, we appeal to the following characterization [D2, Theorem A.1]: a zero dimensional dynamical system (X,T) is asymptotically h-expansive if and only if it is topologically conjugate to a subsystem of a finite entropy countable product of subshifts. It is easy to see that the natural entropy structure of such a product converges uniformly, and this uniform convergence is inherited by systems conjugate to subsystems.

The next theorem is a special case of Ledrappier's result that principal extensions respect the topological conditional entropy defined by Misiurewicz [M2].

Theorem 7.4. [Le] *Let (X',T') be a principal extension of (X,T). Then (X',T') is asymptotically h-expansive if and only if (X,T) is asymptotically h-expansive.*

For the proof of the Sex Entropy Theorem, we only use Theorem 7.4 to know in the proof of Theorem 7.5 below that a zero dimensional principal extension of an asymptotically h-expansive system must be asymptotically h-expansive. Theorem 7.5 shows that passing to principal extensions does not restrict much the variety of symbolic extensions.

Theorem 7.5. *Let (X,T) be an arbitrary topological dynamical system with finite topological entropy. Let $\varphi\colon\ (X',T') \to (X,T)$ be a zero dimensional principal extension. If $\pi\colon\ (Y,S) \to (X,T)$ is a symbolic extension, then (Y,S) has a symbolic principal extension (Y',S) which is an extension of (X',T').*

Proof. The fiber product (Z,S') of Lemma 7.2 is a principal extension of the symbolic (and hence zero dimensional and asymptotically h-expansive) system (Y,S). By Theorem 7.4, (Z,S') is asymptotically h-expansive itself. Z is zero dimensional, as a subspace of the zero dimensional product $Y\times X'$, so it follows from Remarks 7.3 that $\mathcal{H}$ converges uniformly, where $\mathcal{H}$ is defined by 5.2(I) from some refining sequence of clopen partitions. It then follows from Proposition 3.2 that $\mathrm{E}\mathcal{H} = h$, so by Case Z of the Sex Entropy Theorem, (Z,S') admits a principal symbolic extension, which then becomes both an extension of (X',T') and a principal extension of (Y,S), as required. □

Let $P_k = \{P_{k,i}\}$ be a refining sequence of finite small-boundary partitions for a system (X,T) (recall Definitions 5.1). We recall a standard construction, which associates to this sequence a zero dimensional principal extension (X',T') of (X,T) (and which reduces to the construction described in Notation 6.2 when the $P_{k,i}$ are clopen). Given x in X and $k \in \mathbb{N}$, define a sequence $x^k = (x^k_n)_{n\in\mathbb{Z}}$ by setting $x^k_n = i \iff T^n(x) \in P_{k,i}$. Let X'_k be the closure of $\{x^k\colon x \in X\}$. The rule $x^{k+1} \mapsto x^k$ determines a factor map of subshifts $(X'_{k+1},S) \to (X'_k,S)$. These maps, used as bonding maps, define (X',T') as an inverse limit system. The map $\pi\colon (X',T') \to (X,T)$ is the equivariant map which sends the subset of (X',T') with zero coordinate i in the coordinate for X'_k onto the closure of $P_{k,i}$. Any point of X with more than one preimage under π must lie in the boundary of some set $T^j P_{i,k}$. The union of all such boundaries has measure zero for every μ in $K(X,T)$, and therefore the inverse image of this union has measure zero for every μ in $K(X',T')$. It follows that the map π induces an affine homeomorphism $\pi\colon K(X',T') \to K(X,T)$ which for every μ in $K(X',T')$ is a measure theoretic isomorphism between the measure preserving systems (X',μ,S) and $(X,\pi\mu,T)$, and in particular respects measure theoretic entropy, so π is a principal extension. Finally, the entropy structure $\mathcal{H}$ on $K(X,T)$ given by $h_k(\mu) = h(T,\mu,\mathcal{P}_n)$ is simply the renaming by π of the standard entropy structure of the inverse limit system $K(X',T')$, which we will call $\mathcal{H}'$.

With this in hand, it is now easy to finish the proof of Theorem 5.5 in the case that the entropy structure $\mathcal{H}$ on (X,T) is given from a sequence P_k as above.

Proof of Theorem 5.5 in Case I *of* $\mathcal{H}$.

First, if E is a bounded affine superenvelope of $\mathcal{H}$, then $E' := E\circ\pi$ is a bounded affine superenvelope of $\mathcal{H}'$, so by the Case Z of Theorem 5.5 proved in the last section there is a symbolic extension φ of (X,T) such that $h^{\varphi}_{\text{ext}} = E'$, and then $\pi\circ\varphi$ is a symbolic extension of (X,T) such that $h^{\pi\circ\varphi}_{\text{ext}} = E$. Conversely, if $\varphi\colon (Y,S)\to(X,T)$ is a symbolic extension of (X,T), then by Theorem 7.5 there is a symbolic extension $\psi\colon (Y',S)\to(X',T')$ such that $h^{\varphi}_{\text{ext}} = h^{\psi}_{\text{ext}}\circ\pi$. Because h^{ψ}_{ext} is a bounded affine superenvelope of $\mathcal{H}'$, it follows that h^{φ}_{ext} is a bounded affine superenvelope of $\mathcal{H}$. □

Existence of principal extensions

Not every system has a refining sequence of small boundary partitions; for example, an interval of fixed points is an obstruction. In the next theorem, we note that many do. The theorem is essentially a collection of results of Kulesza, Lindenstrauss and Weiss.

Theorem 7.6. *Each of the following conditions is sufficient to guarantee that a system* (X,T) *has a refining sequence of small boundary partitions.*

(1) X *is zero dimensional.*
(2) X *is finite dimensional and the periodic point set of* T *is zero dimensional.*
(3) (X,T) *has an infinite minimal factor and* $\mathbf{h}_{\text{top}}(T) < \infty$.

Remark 7.7. It is the last condition (3) on which we will rely to construct an entropy structure for a general system.

Proof. (1) This is trivial.

(2) This is essentially contained in [Ku], as described in [BFF, App. B]).

(3) In this case the existence of the small boundary partitions follows from two deep results of Lindenstrauss and Weiss involving the *mean dimension* (an idea suggested by Gromov). First, every finite entropy topological dynamical system has zero mean dimension ([LiWe, Theorem 4.2] and the discussion above it); and second, every zero mean dimension dynamical system admitting a nonperiodic minimal factor has the so called *small boundary property* [Li, Theorem 6.2], which implies the existence of a basis of the topology consisting of sets whose boundaries have measure zero for every invariant measure. With these results it is easy to construct the refining sequence of partitions with small boundaries. □

We get as a corollary a very general statement.

Proposition 7.8. *Every finite entropy system* (X,T) *has a zero dimensional principal extension.*

Proof. Let (Z,R) be an infinite, zero entropy minimal system. Then $(X\times Z, T\times R)$ is a principal extension of (X,T). By Fact 7.6(3), $(X\times Z, T\times R)$ has a zero dimensional principal extension, which by composition is a principal extension of (X,T). ☐

In the special case that (X,T) is asymptotically h-expansive, there is a more direct construction of a zero dimensional principal extension [BFF, A.2 and 6.7].

Proof of Theorem 5.5 in Case II *of* $\mathcal{H}$.

We isolate the main argument of the proof as the next lemma.

Lemma 7.9. *Given a system (X,T), let (Z,R) be an aperiodic minimal system with an R-invariant Borel probability measure λ such that $h(R,\lambda)=0=\mathbf{h}_{\rm top}(R)$. Set $(X',T')=(X\times Z, T\times R)$ and let $\pi: X'\to X$ denote the projection on the first axis. Let (α_k) be a refining sequence of partitions on X' with small boundaries for the system (X',T') and let $\mathcal{H}'=(h'_k)$ be the corresponding entropy structure on $K(X',T')$, defined for T' as in* 5.2(I). *Let $\mathcal{H}''=(h''_k)$ be the sequence of functions on $K(X',T')$ defined by $h''_k(\mu')=h'_k(\mu\times\lambda)$, where $\mu=\pi\mu'$.*

Then the following are equivalent.

1. *E is a bounded superenvelope for $\mathcal{H}''$.*
2. *E is a bounded superenvelope for $\mathcal{H}'$.*

Proof. Let h denote the entropy function on $K(X',T')$. Obviously, h is the limit of both $\mathcal{H}'$ and $\mathcal{H}''$.

(1) $\implies$ (2) Let E be a bounded superenvelope of $\mathcal{H}''$. There is no inequality granted between h'_k and h''_k. However, given k, we do have $E-h'_k\geq E-h=\lim_n E-h''_n$, which is nonnegative because E is a superenvelope of $\mathcal{H}''$. It remains to show $E-h'_k$ is u.s.c. Fix $\mu'\in K(X',T')$ and consider the defect

$$\limsup_{\mu'_n\to\mu'}(E-h'_k)(\mu'_n)-(E-h'_k)(\mu').$$

Let D be the compact countable set $\{\mu',\mu'_n,\mu\times\lambda,\mu_n\times\lambda\colon n\in\mathbb{N}\}$, where (μ'_n) is a sequence which realizes the above limsup. We will restrict our considerations to D. We have

$$E-h'_k=(E-h''_\ell)+(h''_\ell-h'_k),$$

for any $\ell>k$. Since the first bracket $(E-h''_\ell)$ is u.s.c., it suffices to estimate the defect of upper semicontinuity of the second bracket (restricted to D) at μ', for which we can choose ℓ as large as we need. Because D is countable, we can find a partition of X, say β, with boundaries of measure zero for all measures arising as projections of measures from D. Also, because α_k has small boundaries, by choosing this partition β fine enough and by choosing a partition γ of Z fine enough, we can approximate the partition α_k by a coarsening β_1 of $\beta\times\gamma$ such that for all measures in D,

$$|h(\cdot,\beta_1)-h(\cdot,\alpha_k)|<\frac{\epsilon}{2}\ .$$

On account of the null boundaries and coarsening, the function $h(\cdot,\beta\times\gamma)-h(\cdot,\beta_1)$ is u.s.c. on D. Further, we can approximate $\beta\times\gamma$ by a coarsening β_2 of a partition α_ℓ (for some large ℓ) so that

$$|h(\cdot,\beta_2)-h(\cdot,\beta\times\gamma)|<\frac{\epsilon}{2}$$

on D. Also, $h(\cdot,\alpha_\ell) - h(\cdot,\beta_2)$ is u.s.c. (this holds not only on D, but even on K). Now observe that $h(\nu',\beta\times\gamma) = h(\nu\times\lambda,\beta\times\gamma)$, because the partition is a product partition and $\mathbf{h}_{\rm top}(R) = 0$. Thus the function $h''_\ell - h'_k$ at a point $\nu' \in K(X',T')$ can be written as

$$\Big[h(\nu\times\lambda,\alpha_\ell) - h(\nu\times\lambda,\beta_2)\Big] \; + \; \Big[h(\nu\times\lambda,\beta_2) - h(\nu\times\lambda,\beta\times\gamma)\Big] \; + \\ \Big[h(\nu',\beta\times\gamma) - h(\nu',\beta_1)\Big] \; + \; \Big[h(\nu',\beta_1) - h(\nu',\alpha_k)\Big] \; .$$

Only the second and fourth brackets contribute to the defect of upper semicontinuity on D, and their joint contribution is less than ϵ.

(2) $\implies$ (1) Let E be a superenvelope of $\mathcal{H}'$. The argument for nonnegativity of $E - h''_k$ is essentially the same as for the first implication. The argument for uppersemicontinuity is also essentially the same, after reversing the roles of ν' and $\nu\times\lambda$:

$$E - h''_k = (E - h'_\ell) + (h'_\ell - h''_k)$$

and

$$(h'_\ell - h''_k)(\nu') \; = \; \Big[h(\nu',\alpha_\ell) - h(\nu',\beta_2)\Big] \; + \; \Big[h(\nu',\beta_2) - h(\nu',\beta\times\gamma)\Big] \; + \\ \Big[h(\nu\times\lambda,\beta\times\gamma) - h(\nu\times\lambda,\beta_1)\Big] \; + \; \Big[h(\nu\times\lambda,\beta_1) - h(\nu\times\lambda,\alpha_k)\Big].$$

□

We now prove the implications of the Sex Entropy Theorem for Case II. We continue the notation of Lemma 7.9.

Proof of Theorem 5.5 (1) $\implies$ (2) *in Case* II *of* $\mathcal{H}$.

Suppose E is a bounded affine superenvelope of $\mathcal{H}$. By Lemma 7.9, the pullback E'' of E to $K(X',T')$ is a bounded affine superenvelope of $\mathcal{H}'$. By the Sex Entropy Theorem in Case I, there is a symbolic extension φ of $(X\times Z, T\times R)$ such that $h^{\varphi}_{\rm ext} = E''$. Then $h^{\pi\circ\varphi}_{\rm ext} = E$.

Proof of Theorem 5.5 (2) $\implies$ (1) *in Case* II *of* $\mathcal{H}$.

Suppose $\varphi\colon (Y,S) \to (X,T)$ is a symbolic extension and $E = h^{\varphi}_{\rm ext}$. By Theorem 7.5, let $\psi\colon (Y',S) \to (X',T')$ be a symbolic extension such that (Y',S) is a principal extension of (Y,S), so $E = h^{\varphi}_{\rm ext} = h^{\pi\circ\psi}_{\rm ext} = (h^{\psi}_{\rm ext})^{[K]}$ where $K = K(X,T)$. Because $h^{\psi}_{\rm ext}$ is a bounded affine superenvelope of $\mathcal{H}'$, by Lemma 7.9 it is also a bounded affine superenvelope of $\mathcal{H}''$, and therefore each $h^{\psi}_{\rm ext} - h''_k$ is u.s.c. It then follows from Remark 2.6 that $(h^{\psi}_{\rm ext} - h''_k)^{[K]}$ is u.s.c. Because h''_k is constant and equal to h_k on fibers of π, we have $(h^{\psi}_{\rm ext} - h''_k)^{[K]} = (h^{\psi}_{\rm ext})^{[K]} - h_k = E - h_k$, so E is a superenvelope of $\mathcal{H}$. As we have seen, it follows from Proposition 2.7 that $E = h^{\pi\circ\psi}_{\rm ext}$ is affine.

□

Remark 7.10. With fiber product constructions, we could have avoided the use of Lemma 7.9 above for the direction (2) $\implies$ (1).

8. Consequences for sex entropy

In this section we collect various (at this point easy) corollaries of our work for sex entropy. Recall our conventions: the infimum of the empty set is $+\infty$, and the only unbounded superenvelope is the constant function $+\infty$.

Theorem 8.1. *The sex entropy function $h_{\rm sex}(T,\cdot)$ is the upper semicontinuous concave function $\mathrm{E}\mathcal{H}$ which is the smallest superenvelope of the entropy structure $\mathcal{H}$. Moreover, there is a variational principle for sex entropy:*

$$\mathbf{h}_{\rm sex}(T) = \mathrm{SUP}(h_{\rm sex}(T,\cdot)) := \sup_{\mu} \inf_{\varphi} \{ h^{\varphi}_{\rm ext}(\mu) \colon \varphi \text{ is a symbolic extension of } T\} \ .$$

Proof. It suffices to consider the case in which $\mathbf{h}_{\rm sex}(T)$ is finite (equivalently, $\mathcal{H}$ has a bounded superenvelope). The first claim is a consequence of the Sex Entropy Theorem and the fact (Proposition 4.3) that $\mathrm{E}\mathcal{H}$ is the infimum of the affine superenvelopes of $\mathcal{H}$. The variational principle then follows from the variational statement in Proposition 4.3. □

Remark 8.2. There is of course no variational principle for residual entropy. Theorem 8.1 implies

$$\mathbf{h}_{\rm res}(T) \le \mathrm{SUP}(h_{\rm res}(T,\cdot)) := \sup_{\mu} \inf_{\varphi} \{ h^{\varphi}_{\rm ext}(\mu) - h(\mu) \colon \varphi \text{ is a symbolic extension of } T\}$$

but it is easy to provide examples in which the inequality is strict. (For example, let (X,T) be the disjoint union of a subshift (X_1,S) and a system (X_2,R) such that $\mathbf{h}_{\rm top}(S) \ge \mathbf{h}_{\rm sex}(R) > \mathbf{h}_{\rm top}(R)$.) This is an example of "sex entropy" being notationally more convenient than "residual entropy".

The characterization of sex entropy as $\mathrm{E}\mathcal{H}$ has considerable content on account of the functional analysis results in sections 2, 3 and 4. The results there for a sequence $\mathcal{F}$ of u.s.c. functions with u.s.c. differences apply to any entropy structure $\mathcal{H} = (h_k)_k$. In particular, the characterizations of superenvelopes by continuous covers (Theorem 2.18) and by the inductive construction of Theorem 3.3 apply. The transfiniteness of the last construction explains for us some of the subtlety of residual entropy.

We apply the mix of functional analysis and the Sex Entropy Theorem in the next theorem, a collection of attainability results.

Theorem 8.3 (Attainability Results). *Suppose (X,T) is a system with entropy structure $\mathcal{H}$ and space of invariant Borel probabilities $K(X,T)$, and (X,T) has some symbolic extension (i.e. $\mathrm{E}\mathcal{H}$ is bounded).*

1. *There is a symbolic extension φ such that $h^{\varphi}_{\rm ext} = h^T_{\rm sex}$ if and only if $\mathrm{E}\mathcal{H}$ is affine. If the ergodic measures are a closed subset of $K(X,T)$, then $\mathrm{E}\mathcal{H}$ is affine.*
2. *There is a symbolic extension (Y,S) of (X,T) such that $\mathbf{h}_{\rm top}(S) = \mathbf{h}_{\rm sex}(T)$ $\iff$ there exists an affine superenvelope $E_{\rm A}$ of $\mathcal{H}$ with $\mathrm{SUP}(E_{\rm A}) = \mathrm{SUP}(\mathrm{E}\mathcal{H})$.*
3. *The sex entropy function $h_{\rm sex}(T,\cdot)$ on $K(X,T)$ achieves its maximum on the closure of the ergodic measures.*

Proof. The characterizations in (1) and (2) follow from the Sex Entropy Theorem 5.5 and the characterization of $h_{\rm sex}(T,\cdot)$ as $\mathrm{E}\mathcal{H}$. The claim in (1) involving ergodic measures follows from the Bauer simplex statement in Theorem 4.6 with $\mathcal{F}$ replaced by $\mathcal{H}$. The claim (3) follows from $h_{\rm sex}(T,\cdot) = \mathrm{E}\mathcal{H}$ and Theorem 4.9. □

Similarly, the results of Sections 2 and 3 for estimating or computing $\mathrm{E}\mathcal{F}$ apply to $\mathrm{E}\mathcal{H}$. Moreover, all abstract candidates $\mathcal{F}$ actually occur as entropy structures:

Theorem 8.4. Entropy Structure Realization Theorem [DS2, Theorem 3] *Let K be a Choquet simplex and let $\mathcal{F} = (f_k)_{k\in\mathbb{N}}$ be a nondecreasing sequence of nonnegative affine u.s.c. functions on K such that each $f_k - f_{k-1}$ is u.s.c. Then there exist a zero dimensional dynamical system (X,T) in the form of an inverse limit of subshifts $(X,T) = \overleftarrow{\lim}_k(X_k,S)$, and an affine homeomorphism $a\colon K(X,T) \to K$, such that for every invariant measure μ on X and every k,*

$$(f_k \circ a)(\mu) = h(\mu_k).$$

Consequently, various entropy examples are immediate consequences of analytic examples, the Sex Entropy Theorem and the Entropy Structure Realization Theorem. Here are some instances of this.

Theorem 8.5 (Nonattainment examples).

(1) *There is a system (X,T) such that for every symbolic extension $\varphi\colon (Y,S) \to (X,T)$, $\mathbf{h}_{\mathrm{top}}(S) > \mathbf{h}_{\mathrm{sex}}(T)$. Moreover, this can happen even if T satisfies $h_{\mathrm{sex}} = \widetilde{h}$.*

(2) *There is a system (X,T) such that $h_{\mathrm{sex}}(T,\mu)$ equals 1 for ergodic μ, but $\mathrm{SUP}(h_{\mathrm{sex}}(T,\cdot)) > 1$.*

Proof. By the Entropy Structure Realization Theorem, there are systems with entropy structures $\mathcal{H}$ matching the sequences $\mathcal{F}$ in Example 4.7 and Example 4.8. Now apply the Sex Entropy Theorem.

□

Similarly the examples 2.19, 2.21 and 2.20 extend to entropy structures on Bauer simplices (see Remark 4.5). We take two relevant superenvelope examples to the level of concrete dynamical systems at the end of this section.

First, though, we want to exhibit a theorem summarizing equivalent conditions for asymptotic h-expansiveness, which is a significant condition for smooth as well as topological dynamics. (A C^∞ system must be asymptotically h-expansive, by Buzzi [Bu], following Yomdin – see [BFF, Theorem 7.8] – but for $r < \infty$, a C^r system need not be asymptotically h-expansive, because upper semicontinuity of the entropy function can fail [M1, N].) The main part of this theorem, the equivalence below of (1) and (2), is known; but we will see how the asymptotically h-expansive systems appear very naturally as a distinguished class in the superenvelope/sex entropy setting.

Theorem 8.6 (Asymptotic h-expansiveness). *The following are equivalent for a system (X,T) with entropy structure $\mathcal{H}$.*

(1) *(X,T) is asymptotically h-expansive.*
(2) *(X,T) has a symbolic extension which is a principal extension.*
(3) *$\mathrm{E}\mathcal{H} = h$.*
(4) *h_k converges uniformly to h.*

Proof. (2) $\iff$ (3) follows from the Sex Entropy Theorem (because h is affine).

(3) $\iff$ (4) follows from Proposition 3.2 with $\mathcal{F} = \mathcal{H}$.

(1) $\implies$ (2) was proved independently in [D2] (in the zero dimensional case) and [BFF] (in general).

(2) $\implies$ (1) follows immediately from Theorem 7.4. □

Remark 8.7. The proof above for equivalence of (1) and (2) left the framework of the current paper. In [D3], a generalized approach to entropy structure gives a very natural proof for the equivalence of (1) and (4), which puts the proof of Theorem 8.6 entirely into the entropy structure setting.

Remark 8.8. Kifer and Weiss [KiWe] have a "relative" theorem analogous to (1) $\Longrightarrow$ (2) above. They introduced *asymptotic entropy expansive random transformations*, and proved a generator theorem allowing their representation by random symbolic subshifts.

Remark 8.9. Let $\mathcal{H}$ be an entropy structure. We have from Theorem 2.18 that $\mathrm{SUP}(\mathrm{E}\mathcal{H})-\mathrm{SUP}(h) = \inf_{\mathcal{G}} \mathrm{SUP}(\Sigma\mathcal{G})-\mathbf{h}_{\mathrm{top}}(T)$, where $\mathcal{G}$ denotes a continuous cover of $\mathcal{H}$. Because $\mathbf{h}_{\mathrm{res}}(T) = \mathbf{h}_{\mathrm{sex}}(T) - \mathbf{h}_{\mathrm{top}}(T) = \mathrm{SUP}(\mathrm{E}\mathcal{H}) - \mathbf{h}_{\mathrm{top}}(T)$, we recover for general X the characterization of $\mathbf{h}_{\mathrm{res}}(X,T)$ given for the zero dimensional case in [D2]: $\mathbf{h}_{\mathrm{res}}(T) = \inf_{\mathcal{G}} \mathrm{SUP}(\Sigma\mathcal{G}) - \mathbf{h}_{\mathrm{top}}(T)$. Note, Theorem 2.18 shows that this infimum is always achieved by some cover $\mathcal{G}$, so the continuous covers alone cannot be adequate to detect the nonattainment phenomenon of Theorem 8.5(1).

We now turn to the promised examples. Other constructions of this kind were presented in [D2], achieving all values for the topological, residual and topological conditional entropies consistent with the known constraints. Quite different methods in [BFF] produced another rich family of examples.

Example 8.10. This is a very simple example with nonzero residual entropy. The strategy is simply to construct a system with $\mathcal{H} = \mathcal{F}$ as in Example 2.19.

Let (X_0, S) denote a strictly ergodic {0-1}-subshift with invariant measure μ_0 and entropy 1. Let B_k ($k \geq 2$) be a sequence of blocks occurring in X_0 of lengths increasing with k and such that $\mu_{B_k} \to \mu_0$. Let X denote the set of all {0-1}-valued matrices $x = (x_{k,n})_{k\in\mathbb{N}, n\in\mathbb{Z}}$ satisfying the following conditions.

(1) The first row x_1 of x either belongs to X_0 or has the form $x_{B_k} = \dots B_k B_k B_k \dots$ for some $k \geq 2$. By taking closure we must also admit that x_1 may have the form $y_1(-\infty, m]y_2[m+1, \infty)$ for some $y_1, y_2 \in X_0$, $m \in \mathbb{Z}$.
(2) If the first row is x_{B_k}, then the kth row of x is an element of X_0.
(3) All other rows are filled with zeros.

The set of matrices so constructed is closed and invariant under the horizontal shift $Sx_{k,n} = x_{k,n+1}$. The structure of invariant measures is as follows: there is one measure $\mu^{(1)}$ supported by matrices having nonperiodic first row and all other rows filled with zeros; this measure is isomorphic to μ_0, its entropy is 1 and $h_k(\mu^{(1)}) = 1$ for $k \geq 1$. Moreover, for each $k > 1$, there are finitely many measures $\mu^{(k,i)}$ supported by matrices having nonzero kth row and periodic first row. Again, each of these measures has entropy 1, with $h_n(\mu^{(k,i)}) = 1$ for $n \geq k$ and $h_n(\mu^{(k,i)}) = 0$ for $n < k$. All measures $\mu^{(k,i)}$ accumulate at $\mu^{(1)}$. The entropy function h is 1 on all measures, so $\mathbf{h}_{\mathrm{top}}(T) = 1$, while $\mathrm{E}\mathcal{H} = 1 + 1^{\mathrm{aff}}_{\{\mu^{(1)}\}}$, hence $\mathbf{h}_{\mathrm{sex}}(X,S) = 2$.

Example 8.11. In this example, $\mathbf{h}_{\mathrm{top}}(S) > \mathbf{h}_{\mathrm{sex}}(T)$ for every symbolic extension S of T. (Moreover, $\mathrm{E}\mathcal{H} = \widetilde{h}$, in particular $\mathbf{h}_{\mathrm{res}}(T) = 0$.) The idea is to obtain a system for which the entropy sequence behaves as in the abstract Example 4.7.

Let X_0 and B_k be as in the previous example. Let C_k ($k \geq 2$) be a block of the following structure:

$$C_k = B_2 \dots\dots\dots\dots\dots\dots\dots B_2 B_3 \dots\dots\dots B_3 \cdots B_k \dots B_k,$$

where the repetitions of B_{i+1} (numerous even for $i = k$) occupy roughly 2^{-i} of the length of C_k, and the precision of these proportions improves with k (the number of repetitions for each i increases with k). It is now seen that the measures μ_{C_k} carried by the periodic orbits of $x_{C_k} = \dots C_k C_k C_k \dots$ converge weakly* to $\sum_{k=1}^{\infty} 2^{-k}\mu_{B_{k+1}}$. Let X be the set of all $\{0\text{-}1\}$-valued matrices $x = (x_{k,n})_{k\in\mathbb{N},n\in\mathbb{Z}}$ satisfying the following conditions.

(1) The first row x_1 of x either belongs to X_0 or it has the form x_{B_k} for some $k \in \mathbb{N}$, or it is x_{C_k} for some $k \geq 2$. By taking closure we must additionaly admit the forms $y_1(-\infty, m]y_2[m+1, \infty)$ with $y_1, y_2 \in X_0$, or $x_{B_k}(-\infty, m]x_{B_{k+1}}[m+1, \infty)$ and also $y_1(-\infty, m]x_{B_2}[m+1, \infty)$.
(2) If the first row is x_{C_k}, then the kth row of x is an element of X_0.
(3) All other rows are filled with zeros.

The structure of invariant measures is now the following: the measure $\mu^{(1)}$ is as in the previous example, and h_1 is the (affine extension of) the characteristic function at $\mu^{(1)}$. This measure is approached by periodic measures $\mu^{(k)}$ supported by matrices with x_{B_k} in the first row (this time there is only one such measure for each $k \geq 2$, so the index i can be dropped; moreover, these measures now have entropy zero). In addition, for each $k \geq 2$, we have finitely many measures $\nu^{(k,j)}$ supported by matrices with x_{C_k} in the first row and a nonperiodic kth row. Then h_k is the (affine extension of) characteristic function of the set of measures $\nu^{(k,j)}$. With increasing k, these latest measures approach the combination $\sum_{k=1}^{\infty} 2^{-k}\mu^{(k+1)}$. The behaviour of $\mathcal{H}$ of Example 4.7 is hence copied.

References

[AE] L. Asimow and A. Ellis, *Convexity Theory and its Applications in Functional Analysis*, London Math. Soc. Monographs **16**, Academic Press (1980).

[BFF] M. Boyle, D. Fiebig and U. Fiebig, *Residual entropy, conditional entropy and subshift covers*, Forum Math. **14** (2002), 713-757.

[BGH] F. Blanchard, E. Glasner and B. Host, *A variation on the variational principle and applications to entropy pairs*, Erg. Th. & Dyn. Syst. **17** (1997), 29-43.

[Bow] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer Lec. Notes in Math. **470**, Springer-Verlag (1975).

[Bu] J. Buzzi, *Intrinsic ergodicity of smooth interval maps*, Israel J. Math. **100** (1997), 125-161.

[C] G. Choquet, *Lectures on Analysis, Vol.* II: *Representation Theory*, W. A. Benjamin, Inc. (1969).

[DGS] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Springer Lec. Notes in Math. **527**, Springer-Verlag (1976).

[D1] T. Downarowicz, *The Choquet simplex of invariant measures for minimal flows* Israel J. Math. **74** (1991), 241-256.

[D2] T. Downarowicz, *Entropy of a symbolic extension of a totally disconnected dynamical system*, Erg. Th. & Dyn. Syst. **21** (2001), 1051-1070.

[D3] T. Downarowicz, *Entropy Structure*, preprint (2003).

[DN] T. Downarowicz and S. Newhouse, *Symbolic extensions in smooth dynamical systems*, preprint (2002).

[DS1] T. Downarowicz and J. Serafin, *Fiber entropy and conditional variational principles in compact non-metrizable spaces*, Fund. Math. **172** (2002), 217-247.

[DS2] T. Downarowicz and J. Serafin, *Possible entropy functions*, Israel J. Math. **135** (2003), 221-251.

[JK] K. Jacobs and M. Keane, *0-1 sequences of Toeplitz type*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 123-131.

[KiWe] Y. Kifer and B. Weiss, *Generating partitions for random transformations*, Erg. Th. & Dyn. Syst. **22** (2002), 1813-1830.

[Ku] J. Kulesza, *Zero-dimensional covers of finite-dimensional dynamical systems*, Erg. Th. & Dyn. Syst. **15** (1995), 939–950.

[Le] F. Ledrappier, *A variational principle for the topological conditional entropy*, Springer Lec. Notes in Math. **729** (1979), Springer-Verlag, 78-88.

[LeWa] F. Ledrappier and P. Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc. **16** (1977), 568-576.

[Li] E. Lindenstrauss, *Mean dimension, small entropy factors and an embedding theorem*, Publ. Math. I.H.E.S. **89** (1999), 227-262.

[LiWe] E. Lindenstrauss and B. Weiss, *Mean topological dimension*, Israel J. Math **115** (2000), 1-24.

[M1] M. Misiurewicz, *Diffeomorphism without any measure with maximal entropy*, Bull. Acad. Polon. Sci., Sér. sci. math., astr. et phys. **21** (1973), 903-910.

[M2] M. Misiurewicz, *Topological conditional entropy*, Studia Math. **55** (1976), 175-200.

[N] S. Newhouse, *Continuity properties of entropy*, Annals of Math. **129** (1989), 215-235.

[O] N. Ormes, *Strong orbit realization for minimal homeomorphisms*, J. Anal. Math. **71** (1997), 103-133.

[Pa] W. Parry, *Entropy and generators in ergodic theory*, W. A. Benjamin, Inc., New York-Amsterdam (1969).

[P1] R. R. Phelps, *Lectures on Choquet's Theorem, second edition*, Springer Lec. Notes in Math. **1757**, Springer-Verlag (2001).

[P2] R. R. Phelps, *Personal communication* (2001).

[T] H. Tong, *Some characterizations of normal and perfectly normal spaces*, Duke Math. J. **19** (1952), 289–292.

[Wa] P. Walters, *An introduction to ergodic theory*, Springer–Verlag (1982).

Department of Mathematics, University of Maryland, College Park, MD 20742-4015, U.S.A.

E-mail address: `mmb@math.umd.edu` *Home page*: `www.math.umd.edu/~mmb`

Institute of Mathematics, Technical University, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

E-mail address: `downarim.pwr.wroc.pl` *Home page*: `www.im.pwr.wroc.pl/~downar`