FORWARD MEAN PROXIMAL PAIRS AND ZERO ENTROPY

T. Downarowicz and Y. Lacroix

November 06, 2010

Abstract. After surveying the known connections between topological entropy zero and nonexistence of certain types of pairs in the system, we supplement them by showing that a topological dynamical system has topological entropy zero if and only if it is a factor of a system with no forward mean asymptotic pairs. This covers two former statements of [O-W] and [D-L].

Acknowledgment

The main result of this paper was predicted in a note attached to the referee report of our former paper [D-L]. We thank the referee for turning our attention to the possibility of such a generalization.

Preliminaries and survey of known results

Let X be a compact metric space with a metric denoted by d. We will consider dynamical systems of the form (X,T), where $T : X \mapsto X$ is continuous, sometimes (when indicated) a homeomorphism. We will be interested in three types of pairs of points in the system. The first notion is classical, the second one was introduced in [O-W], and the last one is our modification generalizing both of them.

Definition 1. A pair of points x, y in X is said to be:

1. *asymptotic*, whenever $\lim_{n\to\infty} d(T^n x, T^n y) = 0$;
2. *mean proximal*, whenever $\lim_{n\to\infty} \frac{1}{2n+1}\sum_{i=-n}^{n} d(T^i x, T^i y) = 0$ (this notion applies to homeomorphisms only);
3. *forward mean proximal*, whenever $\lim_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} d(T^i x, T^i y) = 0$.

Recall that for bounded nonnegative sequences the convergence to zero in the average is equivalent to the convergence to zero along a subsequence of density 1 (this applies as well to sequences indexed by $\mathbb{N}$ as to those indexed by $\mathbb{Z}$).

The condition that a given type of pairs is absent leads to three types of systems. The first one is fairly natural and appears implicitly in several older papers; the notation comes form [D-L], the second one occurs in [O-W], while the last one is, again, our modification that strengthens both preceding cases.

2000 *Mathematics Subject Classification.* 37B10, 37B40.

Key words and phrases. topological entropy, forward mean proximal pair, asymptotic pair.

This paper was written during the first author's visits at ISITV in 2010. Research of the first author is supported from resources for science in years 2009-2012 as research project (grant MENII N N201 394537, Poland)

Definition 2. The system (X,T) is:

1. NAP, if it contains no nontrivial asymptotic pairs;

2. *mean distal* (MD), if it contains no nontrivial mean proximal pairs; (this notion applies to homeomorphisms only).

3. *forward mean distal* (FMD), if it contains no nontrivial forward mean proximal pairs.

Ornstein and Weiss [O-W] mention a weaker version of mean distality, called tightness, which they attribute to H. Furstenberg. Like mean distality, it applies to homeomorphisms only, but it requires fixing an invariant measure μ on the system. The system is *tight* (with respect to μ) if after discarding a set of measure zero there are no nontrivial mean proximal pairs. By choosing a topological model for a measure-theoretic dynamical system, the notion of tightness applies to measure-preserving systems and here it turns out to be an invariant of measure-theoretic isomorphism [O-W, Proposition 3].

We define one more type of pairs and systems, applying to homeomorphisms only. Although these notions turn out "unsuccessful", we include them for completeness of our discussion. We will say that a pair x, y is *bilaterally asymptotic* if both $\lim_{n\to\infty} d(T^n x, T^n y) = 0$ and $\lim_{n\to\infty} d(T^{-n} x, T^{-n} y) = 0$. We say that a system is NBAP if it has no bilaterally asymptotic pairs. We make one more convention. If asymptotic pairs (bilaterally asymptotic pairs, forward mean proximal pairs, or mean proximal pairs) can be eliminated by discarding a set of measure zero for some invariant measure μ, we will say that the system is μ-NAP (μ-NBAP, μ-FMD, μ-MD, respectively). Note that μ-MD means the same as tight.

It is obvious that all asymptotic pairs are forward mean proximal and that bilaterally asymptotic pairs are mean proximal. Thus the following implications hold for topological dynamical systems (assuming where necessary that T is a homeomorphism):

FMD $\implies$ MD and NAP

MD $\implies$ μ-MD and NBAP

NAP $\implies$ NBAP and μ-NAP

On the other hand, there are no direct implications between mean proximality and asymptoticicty of pairs. The former requires examining both the forward and backward orbits (in which aspect it is stronger), but imposes a weaker convergence condition. As a consequence, there are no implications between the properties MD and NAP (or μ-MD and μ-NAP). Appropriate examples can be easily found in the existing literature: the MD extension constructed in [O-W] in the proof of Theorem 6 (rephrased as Theorem 3 below) is (in general) not NAP, the NAP extension constructed in [D-L] to prove Lemma 4.3 (rephrased as Theorem 4 below) is (in general) not MD.

There exist four theorems which connect the above notions with zero entropy.

Theorem 1. [O-W, Theorem 1] *A system (X,T) (with T a homeomorphism) which is μ-MD (i.e. tight) has zero measure-theoretic entropy of μ.*

Theorem 2. [B-H-R] *A system (X,T) which is NAP has zero topological entropy.*

Theorem 3. [O-W, Theorem 6] *An ergodic measure-theoretic dynamical system (X,Σ,μ,T) of entropy zero has a measure-theoretic extension in form of a topological dynamical system (Y,S) (with S a homeomorphism) which is MD.*

Theorem 4. [D-L, Lemma 4.3] *Every topological dynamical system (X, T) of zero topological entropy has a topological extension (Y, S) which is NAP.*

Theorem 2 is in fact proved in [B-H-R] in a stronger version, which can be formulated in a way very similar to Theorem 1 (the passage from Theorem 2a to Theorem 2 is via the variational principle):

Theorem 2a. [B-H-R, Propositions 1 and 3] *If a topological dynamical system (X, T) is μ-NAP for an ergodic measure μ then the measure-theoretic entropy of μ is zero.*

Similarly, the original proof of Theorem 3 gives a slightly different statement: every bilateral subshift of zero topological entropy has a topological extension (Y, S) (which is also a bilateral subshift) which is mean distal. (The actual phrasing of Theorem 3 is then obtained by representing a measure-preserving system of zero entropy as a uniquely ergodic subshift on two symbols.) Because every topological dynamical system of zero topological entropy admits a topological extension in form of a bilateral subshift also of zero entropy (see [B-F-F] [B-D] or [D]), this theorem can be rephrased in a purely topological language, in a way completely analogous to Theorem 4:

Theorem 3a. *Every topological dynamical system (X, T) of zero topological entropy has a topological extension (Y, S) which is MD.*

The situation is pictured on the diagram below. The arrows stand for implications, an arrow with an "*ext*" marks that the property above implies the existence of an extension with the property below. The top six arrows are trivial, the lower four represent Theorems 1, 2a, 3a and 4.

$$
\begin{array}{ccccccccc}
 & & & & \text{FMD} & & & & \\
 & & & \swarrow & & \searrow & & & \\
 & & \text{MD} & & & & \text{NAP} & & \\
 & \swarrow & & \searrow & & \swarrow & & \searrow & \\
\mu\text{-MD} & & & & \text{NBAP} & & & & \mu\text{-NAP} \\
 & \searrow & & & & & & \swarrow & \\
 & & & & h(\mu) = 0 & & & & \\
 & & & & h_{\text{top}} = 0 & & & & \\
 & & & ext \swarrow & & \searrow ext & & & \\
 & & \text{MD} & & & & \text{NAP} & &
\end{array}
$$

It is rather striking, that the properties defined via mean proximality appear on this diagram symmetrically to the ones defined using asymptoticity. We can simplify the diagram and rid it off the measure-theoretic ingredients by noting that any MD system is μ-MD for any invariant measure, which, via Theorem 1 and the variational principle implies zero topological entropy (we can picture an arrow from MD directly to $h_{\text{top}} = 0$). Completely analogous argument yields a symmetric arrow from NAP to zero topological entropy (this is Theorem 2). This symmetry suggests the existence of some natural common generalization of Theorems 1 and

2a (or at least of the implications leading from MD and NAP to $h_{\rm top} = 0$), as well as one for Theorems 3a and 4.

New result

In order to get the first above would-be generalization one seeks for a condition weaker than both MD and NAP. NBAP is a natural candidate here. Does it imply zero topological entropy, or, in a measure-theoretic approach, does μ-NBAP imply $h(\mu) = 0$? Unfortunately, the answer is negative (in any approach) and the path leading via NBAP is a dead end on our diagram. Namely, among so-called *bilaterally deterministic* systems (we skip the definition, as it requires more background; see e.g. [D, Section 3.2]) there are examples which combine two properties: they are NBAP and have positive topological entropy (see e.g. [D, Example 3.2.5]; although it is claimed that the system is μ-NBAP for any invariant measure, it is easy to check that it is in fact NBAP). Having made this constatation we abandon this direction, leaving the question of existence of a common generalization of Theorems 1 and 2a (or their topological analogs) open.

We are much more successful in the other direction: finding a common generalization of Theorems 3a and 4. A natural condition stronger than both MD and NAP is FMD, so it suffices to prove the existence of an FMD (topological) extension for every system with zero topological entropy. And this is exactly what we will do in this paper:

Theorem 5. *Every topological dynamical system (X, T) of zero topological entropy has a topological extension (Y, S) which is FMD.*

Before the proof we make a comment that, unlike in the proof of Theorem 3a, we cannot hope to get the FMD extension in form of a subshift. FMD implies NAP and the only NAP subshifts are those consisting of periodic points only ([B-W]) hence cannot serve as extensions for the entire variety of zero entropy systems. Thus in the proof of Theorem 5 we will have to construct a more abstract extension, in form of an inverse limit of subshifts, similarly as it was done in the proof of Lemma 4.3 in [D-L]. In other details the construction resembles that in the proof of Theorem 6 in [O-W].

Proof of Theorem 5. Let (X, T) be any topological dynamical system with zero topological entropy. We do not assume that T is a homeomorphism or even that it is surjective. It is known (see [D-H] or [D]) that this system can be extended to a zero-dimensional system of zero entropy having an odometer as another factor (and in which T is injective). Furthermore, such a system, being asymptotically h-expansive, admits an extension in form of a bilateral subshift, (that is to say, a forward shift-invariant closed subset of $\Lambda^{\mathbb{Z}}$, Λ a finite set) also having zero topological entropy (see [B-F-F], [B-D] or [D]). Because it has entropy zero, the subshift may be on two symbols, e.g. $\Lambda = \{0, 1\}$. The shift map is not necessarily surjective on the subshift space, but this does not bother us at all. The above discussion allows us to restrict the proof to the case where (X, T) is a bilateral subshift on two symbols, with zero topological entropy, and having an odometer as a topological factor. We are now at the starting point of the proof of Theorem 6 in [O-W] and for some time we will follow nearly identical construction steps.

Let $(p_k)_{k\geq 1}$ be the base of the odometer (i.e., a sequence of positive integers satisfying, for $k \geq 2$, the condition $p_k = q_k p_{k-1}$ where q_k is a natural number larger than 1). Recall that we can always replace such a base by any of its subsequences, this is why later we will have the freedom to assume that the numbers p_k grow as fast as wee need. The odometer factor allows (for each $k \geq 1$) to subdivide every sequence x (an element of X) into blocks of equal lengths p_k in such a way, that every resulting block (we will call it a k-block) is a concatenation of the (k–1)-blocks (precisely, of q_k of them). This parsing (for each parameter k separately) can be determined by examining a finite portion of the sequence x.

We now introduce the following abbreviation:

Definition 3. Two 0-1-valued blocks ξ, ξ' of the same length l are said to be *well separated* if they disagree at at least $\frac{l}{3}$ and at most $\frac{2l}{3}$ positions.

In terms of the Hamming distance d_H, the blocks are well separated if $d_H(\xi, \xi') \in [\frac{1}{3}, \frac{2}{3}]$, or, which is much more practical for us, if simultaneously $d_H(\xi, \xi') \geq \frac{1}{3}$ and $d_H(\bar{\xi}, \xi') \geq \frac{1}{3}$, where $\bar{\xi}$ denotes the negation of ξ (zeros replaced by ones and vice-versa).

In this language, the authors of [O-W] associate injectively to each k-block ω appearing in the system X a 0-1-valued block ξ_ω of the same length as ω (i.e. of length p_k) in such a way that the blocks associated to different blocks ω are not only different but also well separated. Later, when constructing a preimage y of an element x of the system, above a block ω in x they allow in y either ξ_ω or $\bar{\xi}_\omega$. The authors give a very sketchy argument as to how a family of pairwise well separated blocks can be selected and counted. We take the opportunity to present a more detailed description.

A small technical difference between the construction presented here and in [O-W] is that our analog of ξ_ω (although we maintain its length p_k) will depend not only on ω but also on the k-block ω' preceding ω in x. So, formally, we will denote it as $\xi_{\omega'\omega}$, where $\omega'\omega$ is a concatenated pair of k-blocks. The family of blocks $\xi_{\omega'\omega}$, with $\omega'\omega$ ranging over all concatenated pairs of k-blocks appearing in the system, will have the property that any two of them are well separated. Whenever a concatenation $\omega'\omega$ occurs in x, we will allow in the preimage y either $\xi_{\omega'\omega}$ or $\bar{\xi}_{\omega'\omega}$, positioned above the right block ω in x. This technicality will lead to producing preimages of points which differ along an essential subset (of positive upper density) of the positive coordinates; in [O-W] we only knew that they differed along an essential subset of $\mathbb{Z}$, but it could be concentrated on the negative part.

The procedure for $k = 1$ is simpler than for other k. We just select a family $\mathcal{B}_1$ of pairwise well separated 0-1 blocks of length p_1. We are interested in counting the cardinality of $\mathcal{B}_1$ and comparing it with the cardinality of the family $\mathcal{W}_1$ of all (appearing in the system) concatenated pairs $\omega'\omega$ of 1-blocks. Here is how we proceed: The first block in $\mathcal{B}_1$, say B_1, is selected completely arbitrarily (the natural choice here is the block consisting of p_1 zeros). In order to keep the desired separation condition fulfilled we must now eliminate (for the future choices) all blocks B (of length p_1) which are not well separated from B_1. That includes all blocks that differ from B_1 at less than $\frac{p_1}{3}$ positions, (that is, agree with B_1 at more than $\frac{2p_1}{3}$ places) and all blocks that differ from $\bar{B}_1$ at less than $\frac{p_1}{3}$ places. The number of such blocks B can be estimated from above by

$$2 \cdot 2^{p_1 H(\frac{2}{3}, \frac{1}{3})} \cdot 2^{\frac{p_1}{3}},$$

where $H(\frac{2}{3}, \frac{1}{3}) = -\frac{2}{3} \log \frac{2}{3} - \frac{1}{3} \log \frac{1}{3}$, and $2^{p_1 H(\frac{2}{3}, \frac{1}{3})}$ is a well-know estimate for the binomial coefficient counting the number of ways $\frac{2p_1}{3}$ places can be selected out of p_1. The factor $2^{\frac{p_1}{3}}$ in our estimate counts the number of ways the block B can be filled at the other $\frac{p_1}{3}$ places where it is allowed to not coincide with B_1, and the front factor 2 comes from taking into account both B_1 and $\bar{B}_1$. At this point we choose the next member of $\mathcal{B}_1$, say B_2 (of course, avoiding the just eliminated blocks). From now on, in addition to the already eliminated blocks we will have to avoid also the blocks which are not well separated from B_2. The number of these additionally eliminated blocks is estimated by the same expression as before. We can proceed in this manner K times, as long as we have some not eliminated blocks left. This is guaranteed (at least) as long as K satisfies

$$K \cdot 2 \cdot 2^{p_1(H(\frac{2}{3},\frac{1}{3})+\frac{1}{3})} < 2^{p_1}$$

(this is a rather crude estimate, as it does not take into account that the eliminated families will typically largely overlap). Eventually, the family $\mathcal{B}_1$ has cardinality larger than or equal to 2^{sp_1-1}, where $s = \frac{2}{3} - H(\frac{2}{3},\frac{1}{3}))$ (the reader easily verifies that s is positive).

Having done the above, we can proceed to associating the blocks $\xi_{\omega'\omega}$ to the concatenated pairs $\omega'\omega$ of 1-blocks. Since the entropy of our system is zero, choosing p_1 large we can make $\frac{\log \#\mathcal{W}_1}{2p_1}$ as small as we need (recall that $\mathcal{W}_1$ is a subfamily of all blocks of length $2p_1$ appearing in the system). In particular, we can arrange that $\#\mathcal{W}_1 \leq \#\mathcal{B}_1$. Then we define the assignment $\omega'\omega \mapsto \xi_{\omega'\omega}$ to be an arbitrary injective map from $\mathcal{W}_1$ to $\mathcal{B}_1$.

We continue the induction over k. Suppose we have constructed, for some $k > 1$, an injective assignment $\sigma'\sigma \mapsto \xi_{\sigma'\sigma}$ from the collection $\mathcal{W}_{k-1}$ of all concatenated pairs of $(k-1)$-blocks occurring in the system, into the collection of all 0-1-valued blocks of length p_{k-1}, whose range is a pairwise well separated family. We are going to construct the assignment for the parameter k.

We order all concatenated pairs $\omega'\omega$ of k-blocks occurring in the system as $\omega_j'\omega_j$ (j ranges from 1 to the cardinality of all such pairs) and we run an "internal induction" along j. We start with $j = 1$. We will construct $\xi_{\omega_1'\omega_1}$ "above" (i.e., aligned with) ω_1, the right hand part of $\omega_1'\omega_1$. We view ω_1 as a concatenation of q_k (k–1)-blocks, say $\omega_1 = \sigma_1^{(1)}\sigma_1^{(2)}\dots\sigma_1^{(q_k)}$. We also let $\sigma_1^{(0)}$ denote the rightmost (k–1)-block in ω_1'. Following the general scheme, in $\xi_{\omega_1'\omega_1}$, above each $\sigma_1^{(i)}$ ($i = 1,\dots,q_k$) we will allow the choice between only two blocks: $\xi_{\sigma_1^{(i-1)}\,\sigma_1^{(i)}}$ or $\bar{\xi}_{\sigma_1^{(i-1)}\,\sigma_1^{(i)}}$, and in the first step this choice is completely arbitrary, for instance, we can choose $\xi_{\sigma_1^{(i-1)}\,\sigma_1^{(i)}}$ each time. We can now pass to the inductive step j. Looking at the component $(k-1)$-blocks $\sigma_j^{(1)},\dots,\sigma_j^{(q_k)}$ of ω_j and the rightmost component $\sigma_j^{(0)}$ of ω_j' we must decide, for each $i = 1,\dots,q_k$, what to put in ξ_{ω_j',ω_j} above $\sigma_j^{(i)}$: $\xi_{\sigma_j^{(i-1)}\,\sigma_j^{(i)}}$ or $\bar{\xi}_{\sigma_j^{(i-1)}\,\sigma_j^{(i)}}$. These choices (coded by 0's and 1's, respectively) form a binary block (denote it by C_j) of length q_k. We are assuming that the "decision blocks" $C_1,\dots,C_{j-1}$ have already been determined in the preceding steps of the "internal induction" (for instance, C_1 has been decided to be the block of q_k zeros). We must now determine C_j. A priori we have 2^{q_k} choices, but not all of them are satisfactory. For each $j' < j$ we compare the sequences $\sigma_j^{(0)}\sigma_j^{(1)}\dots\sigma_j^{(q_k-1)}$ and $\sigma_{j'}^{(0)}\sigma_{j'}^{(1)}\dots\sigma_{j'}^{(q_k-1)}$ and we mark those indices i for which $\sigma_j^{(i-1)} = \sigma_{j'}^{(i-1)}$. Let r denote the number of marked indices (r depends on j and j', but it is only a temporary notation). What we need is that if $r > 1$ then C_j and $C_{j'}$ restricted to the marked positions are well separated (we will soon explain why). If $r = 0$ or 1 we can ignore this condition (after all, it is then impossible to fulfill it). By the same counting argument as performed in the case $k = 1$, with each j' this eliminates some number of possibilities for the choice of C_j, not more than

$$2 \cdot 2^{rH(\frac{2}{3},\frac{1}{3})} \cdot 2^{\frac{r}{3}}.$$

Since $r \le q_k$ and the elimination is applied $j-1$ times, we conclude that the selection of C_j is possible at least as long as

$$2(j-1) \cdot 2^{q_k(H(\frac{2}{3},\frac{1}{3})+\frac{1}{3})} < 2^{q_k}.$$

We want the procedure to be performed for all j up to the cardinality $\mathcal{W}_k$ of all concatenated pairs of k-blocks appearing in the system. Thus we need the following to be satisfied:

$$\#\mathcal{W}_k \le 2^{sq_k-1}$$

(where s denotes the same positive constant as before). Since the entropy of the system is zero, this condition can be easily achieved by choosing a sufficiently large parameter p_k (given p_{k-1}). In this manner the assignment is completely defined. It remains to verify the separation condition.

Suppose $\omega'\omega$ and $\omega'''\omega''$ are different pairs of k-blocks appearing in the system. They have some labels $j \ne j'$ in the ordering of the family of all such pairs, and we can assume that $j > j'$. By the construction, $\xi_{\omega'\omega}$ is the concatenation of the component blocks $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}}$ (with $i = 1, \dots q_k$), each of them negated or not according to the terms of the "decision block" C_j. The block $\xi_{\omega'''\omega''}$ is built analogously, with the parameter j'. We will compare these blocks restricted to two subsets of coordinates: A, the set of coordinates covered by the components $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}}$ for which $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}} = \xi_{\sigma_{j'}^{(i-1)}\sigma_{j'}^{(i)}}$, and the remaining set B (covered by the components for which this equality fails).

On B we know not only that the covering blocks $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}}$ and $\xi_{\sigma_{j'}^{(i-1)}\sigma_{j'}^{(i)}}$ are different, but also that they are well separated. So, relatively on B, the blocks $\xi_{\omega'\omega}$ and $\xi_{\omega'''\omega''}$ are easily seen to be well separated regardless of the values of C_j and $C_{j'}$.

It remains to prove that the blocks $\xi_{\omega'\omega}$ and $\xi_{\omega'''\omega''}$ restricted to A are well separated. Here $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}} = \xi_{\sigma_{j'}^{(i-1)}\sigma_{j'}^{(i)}}$, hence it is clear that the relative separation depends exclusively on the separation between C_j and $C_{j'}$ restricted to the corresponding indices i. But the equality $\xi_{\sigma_j^{(i-1)}\sigma_j^{(i)}} = \xi_{\sigma_{j'}^{(i-1)}\sigma_{j'}^{(i)}}$ is possible only when $\sigma_j^{(i-1)} = \sigma_{j'}^{(i-1)}$, which happens precisely for the "marked" indices i, and we have arranged that C_j and $C_{j'}$ restricted to the marked indices are well separated. In case $r = 1$ the part A is so small compared to B, that regardless to what happens here, the blocks are well separated (perhaps with a small adjustment of the parameter $\frac{1}{3}$ in the definition of the separation) by being well separated on part B. This concludes the verification of the separation condition.

We can pass to the second (but not last) part of the proof: the construction of a symbolic extension (Y_1, S). This is done analogously as in [O-W], producing an MD extension. As a symbolic system, (Y_1, S) has no chance to be FMD. At this stage we will only have that the factor map collapses all forward mean proximal pairs. We will build an FMD extension in the third part of the proof.

The construction is natural. We allow a 0-1-valued sequence y to belong to the preimage of x if and only if for every k it satisfies the following condition: Above each occurrence of a k-block ω in x, y has either $\xi_{\omega'\omega}$ or $\bar{\xi}_{\omega'\omega}$, where ω' is the k-block preceding this particular occurrence of ω in x. It is elementary to see that so defined set Y_1 of sequences y is nonempty, closed, shift-invariant, and admits precisely two elements above almost every x: some y and its negation $\bar{y}$. This does not apply to points x which map in the odometer to 0 or to any element in the countable orbit of 0. Such points admit a "parsing of infinite order", i.e., a place where for every k two k-blocks meet. It is elementary to see that such points x admit four preimages, say $y'|y$, $\bar{y}'|y$, $y'|\bar{y}$ and $\bar{y}'|\bar{y}$, where y' and y stand for unilateral sequences (over negative and positive integers, respectively), and the vertical bar indicates the infinite parsing in x. It is evident that among such four elements of Y_1 there are two nontrivial asymptotic pairs. So, (Y_1, S) is not FMD. (As we have said, this is inevitable; practically no subshift is FMD.) Note that these asymptotic pairs remain mean distal, as they eventually differ along all negative coordinates. This is why they present no obstacle in the context of Theorem 3. For us they do.

Before we proceed further, we must clarify one small issue. In order to have a well defined factor map from Y_1 to X we must be able to determine the parsing in each element $y \in Y_1$ (i.e., Y_1 must remain an extension of the odometer). Provided this holds, the element x is easily recovered from y by reversing the (injective) assignments $\omega'\omega \mapsto \xi_{\omega'\omega}$. There are many ways of "memorizing" the odometer in the elements y. One possibility is such that when defining the families of pairwise well separated blocks which serve as $\xi_{\omega'\omega}$ we can easily impose the additional requirement, that each of them has a short distinctive block at the left end (a special prefix), which is prohibited to occur otherwise and hence allows to locate the parsing in any concatenation. The cardinality of all 0-1-valued blocks of length p_k with such a distinctive prefix is "almost" 2^{p_k} (on the logarithmic scale), hence such a restriction does affect the construction.

We will now check for forward mean proximal pairs in Y_1. Consider two elements $y_1, y_2 \in Y_1$. If they project to different elements in the odometer, then, since the odometer is equicontinuous, such pair is not proximal (neither forward nor backward), and thus it cannot be forward mean proximal. If y_1 and y_2 project to the same element of the odometer, but to two different points x_1, x_2 in X, then x_1 and x_2 have the same parsing for every k. Since they are different, there is a place n where $x_1(n) \neq x_2(n)$. For every k the coordinate n is covered in x_1 and x_2 by two different (although identically positioned) k-blocks ω_1' and ω_2'. Let ω_1 and ω_2 denote the k-blocks to the right from ω_1' and ω_2' in x_1 and x_2, respectively. Since the pairs $\omega_1'\omega_1$ and $\omega_2'\omega_2$ are different, the blocks $\xi_{\omega_1'\omega_1}$ and $\xi_{\omega_2'\omega_2}$ are well separated, and these are the blocks which occur (perhaps negated) in y_1 and y_2 covering the position $n + p_k$. Because this is true for every k, it is now seen that y_1 and y_2 are not forward proximal (the averages in the definition of forward proximality have upper limit at least $\frac{1}{6}$). We remark, that the simpler construction given in [O-W] gives that the well separated blocks in y_1 and y_2 cover the coordinate n (without "$+p_k$"), admitting the case where they differ only along negative coordinates.

The above discussion proves that any forward proximal pairs y_1, y_2 must map to the same element in X (we have already identified all such pairs as $y'|y$, $\bar{y}'|y$ and $y'|\bar{y}$, $\bar{y}'|\bar{y}$). We have proved that the factor map from Y_1 to X collapses forward mean proximal pairs.

The third and last step of the construction copies verbatim the inverse limit technique from [D-L] to produce an FMP extension of X. In [D-L], we have used it to produce a NAP extension once we knew how to build a symbolic extension which collapses asymptotic pairs. Now we only replace the term "asymptotic" by "forward mean proximal".

Given X, we have already built its symbolic extension Y_1 such that the factor map collapses all forward mean proximal pairs. Notice that Y_1 also has topological entropy zero, because it is a finite-to-one extension of a system of entropy zero. We can now repeat the process and build an extension Y_2 of Y_1 which has topological entropy zero and such that the factor map collapses forward mean proximal pairs. Continuing in this manner we build a sequence of extensions Y_n. In the end we define (Y, S) as the inverse limit system of this sequence. Clearly, it is an extension of (X, T). Suppose it has a distinct forward mean proximal pair. To be distinct in the inverse limit, this pair must be distinct when projected to some Y_n. The projection of this pair to Y_{n+1} must also be distinct, since its elements map to distinct elements in Y_n. But the projection of a forward mean proximal pair is always forward mean proximal (or collapsed). We get that our pair projected to Y_{n+1} is forward mean proximal. But the map from Y_{n+1} to Y_n collapses all such pairs, so the considered pair is not distinct after projecting to Y_n. This contradiction proves that (Y, S) is FMD, concluding the proof. $\square$

We end the paper by formulating (as an obvious consequence of Theorem 5) yet another characterization of systems with zero topological entropy, appending the list given in [D-L].

Corollary. *A topological dynamical system (X, T) has zero topological entropy if and only of it is a factor of an FMD system.*

References

[B-H-R] Blanchard, F., Host, B. and Ruette, S., *Asymptotic pairs in positive-entropy systems*, Ergodic Theory Dynam. Systems **22** (2002), 671–686.

[B-D] Boyle, M. and Downarowicz, T., *The entropy theory of symbolic extensions*, Inventiones Mathematicae **156** (2004), 119–161.

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

[B-W] Bryant, B. and Walters, P., *Asymptotic properties of expansive homeomorphisms*, Math. Systems Theory **3** (1969), 60–66.

[D] Downarowicz, T., *Entropy in Dynamical Systems*, Cambridge University Press, 2011.

[D-H] Downarowicz, T. and Huczek, D., *Zero-dimensional principal extensions*, Acta Applicandae Mathematica (to appear).

[D-L] Downarowicz, T. and Lacroix, Y., *Topological entropy zero and asymptotic pairs*, Israel J. Math. (to appear).

[O-W] Ornstein, D.S. and Weiss, B., *Mean distality and tightness*, Proceedings of the Steklov Institute of Mathematics **244** (2004), 312–319.

Institute of Mathematics and Computer Science, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

E-mail address: downar@pwr.wroc.pl

Institut des Sciences de l'Ingénieur de Toulon et du Var, Laboratoire Sytèmes Navals Complexes, Avenue G. Pompidou, B.P. 56, 83162 La Valette du Var Cedex, France

E-mail address: yves.lacroix@univ-tln.fr