KOLMOGOROV-SINAI ENTROPY VIA SEPARATION PROPERTIES OF ORDER-GENERATED $\sigma$-ALGEBRAS

ALEXANDRA ANTONIOUK
Institute of Mathematics of NAS of Ukraine
Tereshchenkivs’ka str., 3, 01601 Kyiv, Ukraine
email: antoniouk.a@gmail.com

KARSTEN KELLER
Universität zu Lübeck, Institut für Mathematik
Ratzeburger Allee 160, 23562 Lübeck, Germany
email: keller@math.uni-luebeck.de

SERGIY MAKSYMENKO
Institute of Mathematics of NAS of Ukraine
Tereshchenkivs’ka str., 3, 01601 Kyiv, Ukraine
email: maks@imath.kiev.ua

ABSTRACT. In a recent paper, K. Keller has given a characterization of the Kolmogorov-Sinai entropy of a discrete-time measure-preserving dynamical system on the base of an increasing sequence of special partitions. These partitions are constructed from order relations obtained via a given real-valued random vector, which can be interpreted as a collection of observables on the system and is assumed to separate points of it. In the present paper we relax the separation condition in order to generalize the given characterization of Kolmogorov-Sinai entropy, providing a statement on equivalence of $\sigma$-algebras. On its base we show that in the case that a dynamical system is living on an $m$-dimensional smooth manifold and the underlying measure is Lebesgue absolute continuous, the set of smooth random vectors of dimension $n > m$ with given characterization of Kolmogorov-Sinai entropy is large in a certain sense.

1991 Mathematics Subject Classification. Primary: 37A35, 54E52; Secondary: 53C35.

Key words and phrases. Kolmogorov-Sinai entropy, Permutation entropy.

The authors were supported by Marie Curie Actions - International Research Staff Exchange Scheme (IRSES) FP7-People-2011-IRSES. Project number 295164.

The first author was also partially supported by the grant no. 01-01-12 of National Academy of Sciences of Ukraine (under the joint Ukrainian-Russian project of NAS of Ukraine and Russian Foundation of Basic Research).
1. Introduction

1.1. Motivation. Kolmogorov-Sinai entropy of a $\mu$-preserving map $T$ on a probability space $(\Omega, F, \mu)$ is an important concept in dynamical systems and ergodic theory. It is defined as the supremum of the entropy rates $h_\mu(T, A)$ of all finite partitions $A \subset F$ of $\Omega$, which usually makes its determination complicated. In some exceptional cases, a generating partition is known allowing to determine the Kolmogorov-Sinai entropy on the base of only this partition (see, e.g., [18]), but generally one has to take into account an infinite collection of finite partitions.

Here the question arises whether such a collection is given in a natural way. An interesting approach leading to some kind of natural partitioning was given by introducing the concept of permutation entropy by C. Bandt and B. Pompe [5] (see also [2]). This quantity is based on only considering the order structure of a system and has been applied to the analysis of long time series, for example, of electroencephalograms and cardiograms. The point that Kolmogorov-Sinai entropy and permutation entropy coincide for piecewise monotone interval maps, as shown by C. Bandt, G. Keller and B. Pompe [4], gives rise to the question if both entropies are equivalent for a broader class of dynamical systems.

Remark. J. Amigó, M. Kennel, and L. Kocarev [3, 1] have shown equivalence of Kolmogorov-Sinai entropy to a modified concept of permutation entropy which is structurally similar to that of Kolmogorov-Sinai entropy.

K. Keller and M. Sinn [13, 12, 10] have discussed the question of coincidence of permutation entropy and Kolmogorov-Sinai entropy in a general context, in particular by considering dynamical systems equipped with a random vector

$$\Theta = (\xi_1, \xi_2, \ldots, \xi_n) : \Omega \to \mathbb{R}^n.$$ 

Here the idea is to measure complexity of a system via the ‘observables’ $\xi_1, \xi_2, \ldots, \xi_n$. For given $d \in \mathbb{N}$, the set $\Omega$ is partitioned into sets of points $\omega \in \Omega$ for which all vectors

$$(\xi_i(\omega), \xi_i(T(\omega)), \ldots, \xi_i(T^d(\omega))), \quad i = 1, 2, \ldots, d$$

are of the same order type. The larger $d$ the more information on the system is given by the partition obtained in this way and called $P_{\Theta, T}^d$ here. The permutation entropy is defined as the upper limit of the Shannon entropy of the $P_{\Theta, T}^d$ relative to $d$ for $d \to \infty$. 
It has been shown that under certain ‘separation’ conditions on 
\((T, \Theta)\) it holds

\[ h_{\mu}^{KS}(T) = \lim_{d \to \infty} h_{\mu}(T, \mathcal{P}_d^\Theta T), \]

and that the permutation entropy with respect to \(\xi\) is not less than 
the Kolmogorov-Sinai entropy. Under validity of (1), the problem of 
equality of both entropies is reduced to a combinatorial problem related 
to the problem of equality of permutation entropy and the right side of 
(1) (see K. Keller, A. Unakafov and V. Unakafova [11]). Therefore, it 
is of some particular interest to find sufficient conditions for (1) being 
as general as possible. This is the central aim of the present paper.

1.2. An outline. The main ingredient for showing (1) is the equivalence 
of two \(\sigma\)-algebras with respect to \(\mu\) in the case of ergodic \(T\):

\[ \Sigma_{\Theta, T} \overset{\circ}{=} \mathcal{F}, \]

where \(\Sigma_{\Theta, T}\) is the \(\sigma\)-algebra generated by \(\bigcup_{d=1}^{\infty} \mathcal{P}_d^\Theta T\). For making 
apparent the structural arguments, consider the third \(\sigma\)-algebra \(\sigma(\{\Theta \circ T^k\}_{k \geq 0})\), generated by \(\Theta\) and their ‘shifts’ \(\Theta \circ T, \Theta \circ T^2, \ldots\). The central 
statement of this paper is that for ergodic \(T\)

\[ \sigma(\{\Theta \circ T^k\}_{k \geq 0}) \subset \Sigma_{\Theta, T}. \]

Since

\[ \Sigma_{\Theta, T} \subset \sigma(\{\Theta \circ T^k\}_{k \geq 0}), \]

which can be verified by standard arguments, this provides that in the 
ergodic case

\[ \sigma(\{\Theta \circ T^k\}_{k \geq 0}) \overset{\circ}{=} \mathcal{F} \]

is equivalent to (2), hence sufficient for (1). The second ingredient for 
showing (1) is ergodic decomposition.

Condition (5) is substantially weaker than the corresponding statement 
in [10], allowing generalizations of consequences of the main statement 
therein. In particular, the application of embedding theory (compare [16] and [17]) is more apparent from the viewpoint of our paper, 
but it also turns out that the full power of this theory is not needed. In 
this paper we will show that the set of smooth maps \(\Theta\) satisfying (1) 
which are not too far from being injective is large in a certain sense.
1.3. Organization of the paper. In Section 2 we give the basic definitions and formulate the main statements of the paper, which are Theorems 2.5 and 2.13.

Section 3 is mainly devoted to the proof of Theorem 2.5. Here, the ideas given in [10] are lifted to a sufficiently abstract level, in order to extract the general structures and to find the necessary assumptions under which (1) is satisfied.

The proof of Theorem 2.13 is given in Section 4. As a preparation of the proof, we recall some definitions and statements from (differential) topology, as for example the Multijets transversality theorem, and deduce some statement being interesting from their own right.

2. Preliminaries and formulation of main results

2.1. Kolmogorov-Sinai entropy. Let \( \Omega \) be a non-empty set. For a family of subsets \( \mathcal{A} = \{A_i\}_{i \in I} \) of \( \Omega \), denote by \( \sigma(\mathcal{A}) \) the \( \sigma \)-algebra generated by \( \mathcal{A} \).

If \( \Theta : \Omega \to X \) is a map into some topological space \( X \), then we denote by \( \sigma(\Theta) \) the \( \sigma \)-algebra on \( \Omega \) of inverse images of the \( \sigma \)-algebra \( \mathcal{B}(X) \) of Borel subsets of \( X \) under \( \Theta \).

If \( \mathcal{A} = \{A_i\}_{i \in I} \) and \( \mathcal{B} = \{B_j\}_{j \in J} \) are two partitions of \( \Omega \), then we define the new partition \( \mathcal{A} \vee \mathcal{B} \) of \( \Omega \) by
\[
\mathcal{A} \vee \mathcal{B} = \{A_i \cap B_j \mid A_i \in \mathcal{A}, B_j \in \mathcal{B}\}.
\]

We write
\[
\mathcal{A} \prec \mathcal{B}
\]
if each element \( A \in \mathcal{A} \) is a finite union of some elements of \( \mathcal{B} \).

Let \( \mathcal{F} \) be a \( \sigma \)-algebra of subsets of \( \Omega \) and \( \mu \) be a measure on \( \mathcal{F} \). Denote by \( \Pi(\mathcal{F}) \) the set of all finite partitions \( \mathcal{A} = \{A_1, \ldots, A_n\} \) of \( \Omega \) such that \( A_i \in \mathcal{F} \) for each \( i = 1, \ldots, n \). Then the entropy of \( \mathcal{A} \in \Pi(\mathcal{F}) \) with respect to \( \mu \) is defined by the formula
\[
H_\mu(\mathcal{A}) = -\sum_{i=1}^{n} \mu(A_i) \log \mu(A_i).
\]

Further, let \( T : \Omega \to \Omega \) be a measurable map. Denote by \( T^{-1} \mathcal{A} \) the partition of \( \Omega \) consisting of all inverse images of elements of \( \mathcal{A} \):
\[
T^{-1} \mathcal{A} = \{T^{-1}(A_1), \ldots, T^{-1}(A_n)\}.
\]

For each \( k \geq 1 \) define the partition
\[
\tau_k(\mathcal{A}) = \mathcal{A} \vee T^{-1} \mathcal{A} \vee \cdots \vee T^{-(k-1)} \mathcal{A}.
\]

Evidently,
\[
\tau_1(\tau_k(\mathcal{A})) = \tau_{k+1}(\mathcal{A}).
\]
Definition 2.1. Let $T : \Omega \to \Omega$ be a measurable map. Then its Kolmogorov-Sinai entropy is defined by the formula:

$$h^{KS}_\mu(T) = \sup_{A \in \Pi(F)} \lim_{k \to \infty} \frac{1}{k} H_\mu(\tau^k(A)).$$

Though the computation of Kolmogorov-Sinai entropy requires considering all finite partition of $\Omega$ belonging to $\Pi(F)$, the following lemma shows that this entropy can be obtained from certain increasing sequences of finite partitions.

Lemma 2.2. [18, Lemma 4.2] Let $\{A_d\}_{d \geq 1}$ be a sequence of finite partitions of $F$ such that

$$A_1 \prec A_2 \prec \cdots \prec A_d \prec \cdots$$

and $\sigma(\{A_d\}_{d=1}^\infty) = F$. Then $h^{KS}_\mu(T) = \lim_{d \to \infty} \lim_{k \to \infty} \frac{1}{k} H_\mu(\tau^k(A_d)).$

If $A, B \subset F$ are two sub-$\sigma$-algebras, we write $B \circlearrowleft A$ if for each $B \in B$ there exists some $A \in \mathcal{A}$ such that $\mu(B \triangle A) = 0$. Correspondingly, we write $B \circlearrowright A$ if $A \circlearrowleft B$ and $B \circlearrowright A$.

2.2. Ordinal partition $O_d$ of $\mathbb{R}^{d+1}$. For a permutation $\pi = (i_0, \ldots, i_d)$ of a set $\{0, \ldots, d\}$ define the subset $O_\pi$ of $\mathbb{R}^d$ by the following rule: the point $(x_0, \ldots, x_d) \in \mathbb{R}^{d+1}$ belongs to $O_{i_0, \ldots, i_d}$ whenever

$$x_{i_0} \geq x_{i_1} \geq \cdots \geq x_{i_d}$$

and if $x_{i_\tau} = x_{i_{\tau+1}}$ for some $\tau \in \{0, \ldots, d-1\}$ then

$$i_\tau > i_{\tau+1}.$$

Remark. Notice that each vector $x = (x_0, \ldots, x_d) \in \mathbb{R}^{d+1}$ can be regarded as a $(d + 1)$-tuple of pairs of numbers:

$$\left( (x_0, 0), (x_1, 1), \ldots, (x_d, d) \right).$$

This set can be uniquely lexicographically ordered in a decreasing manner: at first we sort them by values of $x_i$, and then by their indices $i$. Thus we can associate to $x$ a unique permutation $\pi$ of indexes $\{0, \ldots, d\}$ which sorts the above set of pairs (6). Then $O_\pi$ consists of all $x \in \mathbb{R}^{d+1}$ that can be sorted by the same permutation $\pi$.

It is easy to see that the following family of sets

$$O_d = \{O_\pi \mid \pi = (i_0, \ldots, i_d) \text{ is a permutation of } \{0, \ldots, d\} \}$$

is a partition of $\mathbb{R}^{d+1}$.
2.3. **Ordinal partition of** $\Omega$. Now let $\Omega$ be a set, $T : \Omega \to \Omega$ and $\xi : \Omega \to \mathbb{R}$ be a function. Then for each $d \in \mathbb{N}$ we can define the following map

$$\Lambda_d = (\xi, \xi \circ T, \ldots, \xi \circ T^d) : \Omega \to \mathbb{R}^{d+1}.$$ 

Define the partition $\mathcal{P}^\xi_T = \{ P^\xi_{\pi, T} \}$ of $\Omega$, where

$$P^\xi_{\pi, T} = \Lambda_d^{-1}(O_\pi),$$

and $\pi$ runs over all permutations of the set $\{0, \ldots, d \}$. Thus $\mathcal{P}^\xi_T$ is just the inverse image of the partition $O_d$ of $\mathbb{R}^{d+1}$ under the map $\Lambda_d$.

**Remark.** Notice that each set $P^\xi_{\pi, T}$, $\pi = (i_0, \ldots, i_d)$ consists of all $\omega \in \Omega$ such that

$$\xi \circ T^{i_0}(\omega) \geq \xi \circ T^{i_1}(\omega) \geq \cdots \geq \xi \circ T^{i_d}(\omega),$$

and if $\xi \circ T^{i_\tau}(\omega) = \xi \circ T^{i_{\tau+1}}(\omega)$, then $i_\tau > i_{\tau+1}$.

**Remark.** The partition $\mathcal{P}^\xi_T$ can also be described in the following way. For each pair $(i, j)$ such that $0 \leq i < j \leq d$ define the partition of $\Omega$ by two sets:

$$R^{i,j}_d = \{ \omega \in \Omega \mid \xi \circ T^i(\omega) < \xi \circ T^j(\omega) \},$$

$$R^{j,i}_d = \{ \omega \in \Omega \mid \xi \circ T^i(\omega) \geq \xi \circ T^j(\omega) \}.$$ 

Then it is easy to see that

$$\mathcal{P}^\xi_T = \bigvee_{i \neq j \in \{0, \ldots, d \}} R^{i,j}_d.$$ 

**Definition 2.3.** The $\sigma$-algebra

$$\Sigma^\xi_T = \sigma(\{ \mathcal{P}^\xi_{\pi, T} \})$$

is called the **ordinal** $\sigma$-algebra of $\Omega$ for $(\xi, T)$.

More generally, let $\Theta = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbb{R}^n$ be a map. Then we define the partition

$$P^\Theta_T = \bigvee_{i=1}^n P^\xi_{i, T}, \quad d \geq 1,$$

and the $\sigma$-algebra

$$\Sigma^{\Theta, T} := \sigma\left( \{ P^\Theta_{d, T} \} \right) = \sigma\left( \{ \Sigma^\xi_{i, T} \} \right),$$

which we call the **ordinal** $\sigma$-algebra of $\Omega$ for $(\Theta, T)$. 

Suppose $F$ is a $\sigma$-algebra of subsets of $\Omega$ such that $T : \Omega \to \Omega$ is $F$-$F$-measurable and $\Theta : \Omega \to \mathbb{R}^n$ is $F \cdot \mathcal{B}(\mathbb{R}^n)$-measurable. Then it is obvious that

$$\Sigma_{\Theta, T} \subset F.$$  

The following lemma easily follows from (7) and (8) and we left it to the reader.

**Lemma 2.4.** $\Sigma_{\Theta, T} \subset \sigma(\Theta \circ T_k)_{k \geq 0}$.

### 2.4. Main results.

The following theorem gives sufficient conditions for the validity of (1).

**Theorem 2.5.** Let $(\Omega, F, \mu)$ be a probability space, $T : \Omega \to \Omega$ be a measurable $\mu$-invariant transformation, and $\Theta = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbb{R}^n$ be a measurable map such that $\sigma(\Theta \circ T_k)_{k \geq 0} = F$. Suppose also that one of the following conditions holds true: either

(a) $T$ is ergodic, or

(b) $T$ is not ergodic, however $\Omega$ can be embedded into some compact metrizable space so that $F = \mathcal{B}(\Omega)$.

Then

$$h_{\mu}^{KS}(T) = \lim_{d \to \infty} \lim_{k \to \infty} \frac{1}{k} H_{\mu} \left( \tau_k(P_{d, T}^{\Theta}) \right).$$

We will now recall the notions of residuality and prevalence being respectively a topological and a measure-theoretic formalization of the expression “almost every”.

A subset $A$ of a topological space is residual if $A$ is an intersection of countably many sets with dense interiors. A Baire space is a topological space in which every residual subset is dense. Every complete metric space is Baire.

**Definition 2.6.** Let $V$ be a linear vector space over $\mathbb{R}$ or $\mathbb{C}$. A finite-dimensional subspace $P \subset V$ is called a probe for a set $A \subset V$ if for each $v \in V$ the intersection $P \cap [(V \setminus A) + v]$ has Lebesgue measure zero in $P$.

Suppose now that $V$ is a topological vector space, i.e. that it has a topology in which addition of vectors and multiplication by scalars are continuous operations.

**Definition 2.7.** Let $\mu$ be a nonnegative measure on the $\sigma$-algebra $\mathcal{B}(V)$, and $S \subset V$ be a Borel subset. Then $\mu$ is said to be transverse to $S$ if the following two conditions hold:

(i) There exists a compact subset $U \subset V$ such that $0 < \mu(U) < \infty$. 

(ii) \( \mu(S + v) = 0 \) for every \( v \in V \), where \( S + v = \{ s + v \mid s \in S \} \) is the translation of \( S \) by the vector \( v \).

**Definition 2.8.** A subset \( A \subset V \) is called prevalent if its complement \( V \setminus A \) is contained in some Borel set admitting a transverse measure.

The following lemma summarizes some properties of prevalent sets obtained in [9].

**Lemma 2.9.** [9] Suppose \( V \) admits a complete metric. Let also \( A \subset V \) be a subset.

1. If \( A \) is prevalent, then \( V \setminus A \) is nowhere dense.
2. If \( \dim V < \infty \), then \( A \) is prevalent if and only if \( V \setminus A \) has Lebesgue measure zero.
3. If \( \dim V = \infty \) and \( V \setminus A \) is compact, then \( A \) is prevalent.
4. If \( A \) admits a probe, then \( A \) is prevalent.

In general classes of residual and prevalent subsets of a complete metric space \( V \) are distinct and no one of them contains the other.

Let \( \Omega \) and \( X \) be smooth manifolds and \( r = 0, 1, \ldots, \infty \). Then the space \( C^r(\Omega, X) \) admits two natural topologies weak, \( C^k_W \), and strong, \( C^k_S \). The following lemma collects some information about these topologies, see e.g. [8, Chapter 2] and [6, Chapter II, §3].

**Lemma 2.10.** 1) Topology \( C^r_S \) is finer than \( C^r_W \). If \( \Omega \) is compact, then these topologies coincide.
2) \( C^r(\Omega, X) \) is a Baire space with respect to each of the topologies \( C^r_W \) and \( C^r_S \).
3) \( C^\infty(\Omega, X) \) admits a complete metric with respect to the weak topology \( C^\infty_W \).
4) Suppose \( X = \mathbb{R}^n \), so the space \( C^r(\Omega, \mathbb{R}^n) \) has a natural structure of a linear space. Then \( C^r(\Omega, \mathbb{R}^n) \) is a topological vector space with respect to the weak topology \( C^\infty_W \). However, if \( \Omega \) is non-compact, then \( C^r(\Omega, \mathbb{R}^n) \) is not a topological vector space with respect to the strong topology \( C^r_S \), since the multiplication by scalars is not continuous.

Again let \( \Omega \) be a smooth manifold of dimension \( m \).

**Definition 2.11.** A subset \( D \subset \Omega \) has measure zero if for any local chart \( (U, \varphi) \) on \( \Omega \), where \( U \subset \Omega \) is an open subset and \( \varphi : U \to \mathbb{R}^m \) is a smooth embedding, the set \( \varphi(D \cap U) \) has Lebesgue measure zero in \( \mathbb{R}^m \).

**Definition 2.12.** Let \( \mu \) be a measure on \( \mathcal{B}(\Omega) \). We will say that \( \mu \) is Lebesgue absolute continuous if \( \mu(D) = 0 \) for any subset \( D \subset \Omega \) of measure zero in the sense of Definition 2.11.
Remark. We can reformulate Definition 2.12 as follows. Let $\lambda$ be a Lebesgue measure on $\mathbb{R}^m$ and $(U, \varphi)$ be a local chart on $\Omega$. Since $\varphi$ is an embedding, we can define the induced measure $\varphi_\ast(\lambda)$ on $\mathcal{B}(U)$ by $\varphi_\ast(\lambda)(A) = \lambda(A)$ for all $A \in \mathcal{B}(U)$. Then $\mu$ is Lebesgue absolute continuous if for any local chart $(U, \varphi)$ the restriction of $\mu$ to $\mathcal{B}(U)$ is absolute continuous with respect to $\varphi_\ast(\lambda)$.

Our second result shows that the set of maps $\Theta$ for which (1) holds is "large".

Theorem 2.13. Let $\Omega$ be a smooth manifold of dimension $m$, $\mu$ be a measure on $\mathcal{B}(\Omega)$, $T : \Omega \to \Omega$ be a measurable $\mu$-invariant transformation. Suppose $\mu$ is Lebesgue absolute continuous in the sense of Definition 2.12. Let $V$ be the set of all maps $\Theta \in C^\infty(\Omega, \mathbb{R}^n)$ for which

\begin{equation}
H^K_S(\mu)(T) = \lim_{d \to \infty} \lim_{k \to \infty} \frac{1}{k} H_{\mu}(\tau_k(P^\Theta_d, T))
\end{equation}

holds. If $n > m$, then $V$ is residual in $C^\infty(\Omega, \mathbb{R}^n)$ with respect to strong topology $C^\infty_S$, and prevalent with respect to the weak topology $C^\infty_W$.

3. Separation via $\sigma$-algebras

3.1. Properties of distribution functions. Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and $\xi : \Omega \to \mathbb{R}$ be a measurable function. Let also $F : \mathbb{R} \to [0, 1]$ be the distribution function of $\xi$, i.e.

\[ F(a) = \mu\{\omega \mid \xi(\omega) \leq a\} = \mu(\xi^{-1}(-\infty, a]) \]

It is well-known that $F$ is non-decreasing, right continuous, and that

\[ \lim_{a \to -\infty} F(a) = 0, \quad \lim_{a \to +\infty} F(a) = 1. \]

The latter justifies that $F$ can also be considered as a function from $[-\infty, +\infty]$ into $[0, 1]$.

For further considerations it will be convenient to keep in mind the following commutative diagram:

\begin{equation}
\begin{array}{ccc}
\Omega & \xrightarrow{\xi} & \mathbb{R} \\
F \circ \xi \downarrow & & \downarrow F \\
[0, 1] & & 
\end{array}
\end{equation}

which implies that $F \circ \xi$ is $\mathcal{F}$-$\mathcal{B}(\mathbb{R})$-measurable and so it holds $\sigma(F \circ \xi) \subset \mathcal{F}$.

For each $a \in \mathbb{R}$ define the following two elements of $[-\infty, \infty]$:

\[ a_* = \inf(F^{-1}F(a)), \quad a^* = \sup(F^{-1}F(a)). \]
Moreover, let
\[ C_a = (-\infty, a). \]

**Lemma 3.1.** Let \( a \in \mathbb{R} \). Then the following statements hold true.

1. \( F^{-1}F(a) \) coincides either with \([a_*, a^*] \) or with \([-\infty, a^*] \).
2. \( F^{-1}F(C_{a_*}) = C_{a_*} \).
3. If \( a_* < a \), then
   \[ F^{-1}F(C_a) = C_{a_*} \cup F^{-1}F(a), \]
   so \( F^{-1}F(C_a) \) either equals \((-\infty, a^*) \) or \((-\infty, a^*] \). Moreover, let
   \[ Z = F^{-1}F(C_a) \setminus C_a. \]

Then \( \mu(\xi^{-1}(Z)) = 0 \).

**Proof.**

(1) Evidently, \( F^{-1}F(a) \subset [a_*, a^*] \).

We will now prove that \([a_*, a^*] \subset F^{-1}F(a) \). By definition of infimum and supremum of the set \( F^{-1}F(a) \) there are two sequences \( \{x_i\}, \{y_i\} \subset F^{-1}F(a) \) such that

\[ a_* \leq x_{i+1} \leq x_i \leq \cdots \leq x_1 \leq y_1 \leq \cdots \leq y_i \leq y_{i+1} \leq \cdots \leq a^*, \]

\[ \lim_{i \to \infty} x_i = a_* \text{ and } \lim_{i \to \infty} y_i = a^*. \]

Since \( F \) is nondecreasing and \( F(x_i) = F(y_i) = F(a) \) for all \( i \), it follows that \( F \) is constant on each segment \([x_i, y_i]\), and so

\[ (a_*, a^*) = \bigcup_i [x_i, y_i] \subset F^{-1}F(a). \]

Moreover, from right-continuity of \( F \) we obtain that \( F(a_*) = \lim_{i \to \infty} F(x_i) = F(a) \), hence \([a_*, a^*] \subset F^{-1}F(a) \).

(2) The inclusion \( C_{a_*} \subset F^{-1}F(C_{a_*}) \) is evident. Suppose that there exists some

\[ t \in F^{-1}F(C_{a_*}) \setminus C_{a_*}. \]

This means that

(i) \( t \geq a_* \), and
(ii) \( F(t) \in F(C_{a_*}) \), i.e. \( F(t) = F(s) \) for some \( s < a_* \).

Thus \( s < a_* \leq t \). Since \( F \) is non-decreasing,

\[ F(s) = F(a_*) \overset{(i)}{=} F(a) = F(t), \]

that is \( s \in F^{-1}F(a) \), and therefore \( a_* \leq s \), contradicting the assumption. Thus \( F^{-1}F(C_{a_*}) = C_{a_*} \).

(3) Since \( a_* < a \), \( F(a_*) = F(a) \), and \( F \) is non-decreasing, it follows that

\[ F(C_a) = F((-\infty, a_*)) \cup F([a_*, a)) = F(C_{a_*}) \cup \{F(a)\}, \]
hence
\[ F^{-1}F(C_a) = F^{-1}(F(C_a) \cup \{F(a)\}) \]
\[ = F^{-1}F(C_a) \cup F^{-1}F(a) \overset{(2)}{=} C_a \cup F^{-1}F(a). \]

It follows that \( Z \) either equals \([a, a^*]\) or \([a, a^*]\). Suppose \( Z = [a, a^*] \), then
\[ \mu(\xi^{-1}(Z)) = \mu(\xi^{-1}[a, a^*]) = \mu(\xi^{-1}(-\infty, a^*)) - \mu(\xi^{-1}(-\infty, a)) \]
\[ = F(a^*) - \lim_{t \to a^+} \mu(\xi^{-1}(-\infty, t)) \]
\[ = F(a) - \lim_{t \to a} F(t) = F(a) - \lim_{t \to a} F(a) = 0. \]

Now, let \( Z = [a, a^*] \). Then similarly,
\[ \mu(\xi^{-1}(Z)) = \mu(\xi^{-1}[a, a^*]) = \mu(\xi^{-1}(-\infty, a^*)) - \mu(\xi^{-1}(-\infty, a)) \]
\[ = \lim_{s \to a^+} \mu(\xi^{-1}(-\infty, s)) - \lim_{t \to a} \mu(\xi^{-1}(-\infty, t)) \]
\[ = \lim_{s \to a^+} F(s) - \lim_{t \to a} F(t) = \lim_{s \to a^+} F(a) - \lim_{t \to a} F(a) = 0. \]

Lemma is completed. \( \square \)

**Lemma 3.2.** \( \sigma(F \circ \xi) \overset{\circ}{=} \sigma(\xi) \).

**Proof.** It is easy to see that \( \sigma(F \circ \xi) \subset \sigma(\xi) \). Indeed, let \( A \in \sigma(F \circ \xi) \), so
\[ A = (F \circ \xi)^{-1}(B) = \xi^{-1}F^{-1}(B) \]
for some \( B \in \mathcal{B}([0,1]) \). But \( F^{-1}(B) \in \mathcal{B}(\mathbb{R}) \), hence \( A \in \sigma(\xi) \).

Now we will show that \( \sigma(F \circ \xi) \supset \sigma(\xi) \). For each \( a \in \mathbb{R} \) let
\[ P_a = \xi^{-1}(C_a). \]

Then \( \sigma(\xi) \) is generated by the sets \( P_a \), so it suffices to prove that for each \( a \in \mathbb{R} \) there exists some \( Q_a \in \sigma(F \circ \xi) \) such that \( \mu(Q_a \triangle P_a) = 0 \).

In fact we will put
\[ Q_a = \xi^{-1}F^{-1}F(C_a) = (F \circ \xi)^{-1}F(C_a). \]

Since \( F \) is non-decreasing \( F(C_a) \) is a Borel subset of \([0,1]\), hence \( Q_a \in \sigma(F \circ \xi) \). So it remains to show that \( \mu(Q_a \triangle P_a) = 0 \) for each \( a \in \mathbb{R} \).

First suppose \( a = a^* \). Then by (2) of Lemma 3.1
\[ P_a = P_{a^*} = \xi^{-1}(C_{a^*}) \overset{(2)}{=} \xi^{-1}F^{-1}F(C_{a^*}) = Q_{a^*} = Q_a, \]
hence \( Q_a \triangle P_a = \emptyset \), and so \( \mu(Q_a \triangle P_a) = 0 \).
Now suppose $a_* < a$. Then for $Z = F^{-1}F(C_a) \setminus C_a$ it holds
\[ Q_a = \xi^{-1}F^{-1}F(C_a) = \xi^{-1}(C_a) \cup \xi^{-1}(Z) = P_a \cup \xi^{-1}(Z), \]
Therefore, by (4)
\[ \mu(Q_a \triangle P_a) = \mu(Q_a \setminus P_a) = \mu(\xi^{-1}(Z)) = 0. \]
The Lemma is proved. \qed

3.2. Ergodic properties. Let $T : \Omega \to \Omega$ be a measurable map. Define the function $I_d : \Omega \to \mathbb{R}$ by
\[ I_d(\omega) = \# \{ r = 1, \ldots, d - 1 \mid \xi \circ T^r(\omega) \leq \xi(\omega) \}. \]
So $I_d(\omega)$ is the number of points among the first $d - 1$ points of the $T$-orbit of $\omega$ at which $\xi$ takes values not greater than $\xi(\omega)$.

**Lemma 3.3.** If $T$ is ergodic and $\mu$-preserving, then

(11) \[ F \circ \xi(\omega) = \lim_{d \to \infty} \frac{I_d(\omega)}{d} \]

**Proof.** For each $a \in \mathbb{R}$ consider the following set
\[ K_a = \xi^{-1}(-\infty, a]. \]
Then by definition
\[ F(a) = \mu(K_a) = \mu(\omega \in \Omega \mid \xi(\omega) \leq a). \]
Moreover, as $T$ is ergodic, it follows from Birkhoff’s Ergodic Theorem that there exists a subset $\Omega_a \subset \Omega$ such that $\mu(\Omega_a) = 1$, and for each $\bar{\omega} \in \Omega_a$
\[ \mu(K_a) = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid T^r(\bar{\omega}) \in K_a \} = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq a \}. \]
Take any countable dense subset $S \subset \mathbb{R}$ containing all points of discontinuity of $F$ and let
\[ \bar{\Omega} = \bigcap_{a \in S} \Omega_a. \]
Then $\mu(\bar{\Omega}) = 1$ as well, and for each $a \in S$ and $\bar{\omega} \in \bar{\Omega}$
\[ \mu(K_a) = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq a \}. \]
In particular, if \( \bar{\omega} \in \bar{\Omega} \) is such that \( a = \xi(\bar{\omega}) \in S \), then
\[
F(\xi(\bar{\omega})) = F(a) = \mu(K_a) = \mu(K_{\xi(\bar{\omega})}) = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq \xi(\bar{\omega}) \} = \lim_{d \to \infty} \frac{I_d(\bar{\omega})}{d}.
\]
Thus, \( \{I_d\} \) converges to \( F \circ \xi \) on the set \( \xi^{-1}(S) \cap \bar{\Omega} \). We will prove that in fact this sequence converges to \( F \circ \xi \) on \( \bar{\Omega} \).

Let \( \bar{\omega} \in \bar{\Omega} \) be such that \( a = \xi(\bar{\omega}) \in \mathbb{R} \setminus S \). Then \( F \) is continuous at \( a \).

Choose two sequences \( \{b_i\}_{i \in \mathbb{N}} \subset S \cap (-\infty, a), \quad \{c_i\}_{i \in \mathbb{N}} \subset S \cap (a, +\infty) \) converging to \( a \). Then by construction of \( \bar{\Omega} \) for each \( i \in \mathbb{N} \) we have that
\[
F(b_i) = \mu(K_{b_i}) = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq b_i \},
\]
\[
F(c_i) = \mu(K_{c_i}) = \lim_{d \to \infty} \frac{1}{d} \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq c_i \}.
\]
Since \( b_i < a < c_i \), we see that
\[
\# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq b_i \} \leq I_d(\bar{\omega}) \leq \# \{ r < d \mid \xi \circ T^r(\bar{\omega}) \leq c_i \}.
\]
Hence
\[
F(a) = \lim_{i \to \infty} F(b_i) \leq \lim_{i \to \infty} I_d(\bar{\omega}) \leq \lim_{i \to \infty} I_d(\bar{\omega}) \leq \lim_{i \to \infty} F(c_i) = F(a).
\]
Thus \( \lim_{d \to \infty} I_d(\bar{\omega}) \) exists and coincides with \( F(a) = F(\xi(\bar{\omega})) \), which proves the lemma.

**Corollary 3.4.** Let \( (\Omega, \mathcal{F}, \mu) \) be a probability space and \( T : \Omega \to \Omega \) and
\[
\Theta = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbb{R}^n
\]
be measurable maps. If \( T \) is ergodic and \( \mu \)-preserving, then
\[
\sigma(\Theta) \circ \subset \Sigma^{\Theta, T}.
\]

**Proof.** Suppose \( n = 1 \), so \( \Theta = \xi : \Omega \to \mathbb{R} \) is a function. Then by Lemma 3.2 \( \sigma(\xi) \circ \subset \sigma(F \circ \xi) \). Notice that \( I_d \) is \( \Sigma^{\xi, T} \)-\( \mathcal{B}([0, 1]) \)-measurable for each \( d \) and by Lemma 3.3 the sequence \( \{I_d\} \) converges a.e. to \( F \circ \xi \). Hence \( F \circ \xi \) is \( \mathcal{F} \)-\( \mathcal{B}([0, 1]) \)-measurable as well. This means that \( \sigma(\xi) \circ \subset \sigma(F \circ \xi) \circ \subset \Sigma^{\xi, T} \).

If \( n \geq 2 \), then for each \( i = 1, \ldots, n \) we have the inclusion of \( \sigma \)-algebras:
\[
\sigma(\xi_i) \subset \Sigma^{\xi_i, T}.
\]
Corollary 3.5. Let \((\Omega, \mathcal{F}, \mu)\) be a probability space, \(T : \Omega \to \Omega\) and \(\Theta = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbb{R}^n\) be measurable maps. If \(T\) is ergodic and \(\mu\)-preserving, then

\[
\sigma\left(\{\Theta \circ T^k\}_{k \geq 0}\right) \subset \Sigma^{\Theta, T}.
\]

Proof. Since \(\Sigma^{\Theta, T} \subset \sigma\left(\{\Theta \circ T^k\}_{k \geq 0}\right)\), see Lemma 2.4, it suffices to show that

\[
\sigma\left(\{\Theta \circ T^k\}_{k \geq 0}\right) \subset \Sigma^{\Theta, T}.
\]

As in the proof of Corollary 3.4, we can restrict to the case \(n = 1\) with \(\Theta = \xi : \Omega \to \mathbb{R}\), because \(\Sigma^{\Theta, T}\) is generated by \(\Sigma^{\xi_i, T}, i = 1, \ldots, n\).

We will show that

\[
\Sigma^{\xi \circ T^k, T} \subset \Sigma^{\xi, T}, \quad k \geq 1.
\]

Then by Corollary 3.4 we get the inclusions

\[
\sigma(\xi \circ T^k) \subset \Sigma^{\xi \circ T^k, T} \subset \Sigma^{\xi, T},
\]

which imply \((13)\) with \(\Theta = \xi : \Omega \to \mathbb{R}\).

For the proof of \((14)\) it is sufficient to show that partition \(P_\pi^{\xi \circ T, T}\) is finer than \(P_d^{\xi \circ T, T}\) for all \(d \geq 1:\)

\[
P_d^{\xi \circ T, T} \prec P_\pi^{\xi \circ T, T}, \quad d \geq 1.
\]

Let \(\pi = (i_0, \ldots, i_d)\) be a permutation of the set \{0, \ldots, d\} and \(P_\pi^{\xi \circ T, T}\) be the corresponding element of partition \(P_\pi^{\xi \circ T, T}\), so \(P_d^{\xi \circ T, T}\) consists of all \(\omega \in \Omega\) such that

\[
\xi \circ T \circ T^{i_0}(\omega) \geq \xi \circ T \circ T^{i_1}(\omega) \geq \cdots \geq \xi \circ T \circ T^{i_d}(\omega),
\]

and if \(\xi \circ T \circ T^{i_\tau}(\omega) = \xi \circ T \circ T^{i_{\tau + 1}}(\omega)\), then \(i_\tau > i_{\tau + 1}\).

In other words, \(\omega \in P_\pi^{\xi \circ T, T}\) if and only if

\[
\xi \circ T^{i_0 + 1}(\omega) \geq \xi \circ T^{i_1 + 1}(\omega) \geq \cdots \geq \xi \circ T^{i_d + 1}(\omega),
\]

and whenever \(\xi \circ T^{i_\tau + 1}(\omega) = \xi \circ T^{i_{\tau + 1} + 1}(\omega)\), then \(i_\tau > i_{\tau + 1}\) for \(\tau \in \{0, \ldots, d - 1\}\).

Consider the following permutations of the set \{0, \ldots, d + 1\}:

\[
\alpha_0 = (0, i_0 + 1, i_1 + 1, \ldots, i_d + 1),
\]

\[
\alpha_1 = (i_0 + 1, 0, i_1 + 1, \ldots, i_d + 1),
\]

\[
\cdots \cdots
\]

\[
\alpha_d + 1 = (i_0 + 1, i_1 + 1, \ldots, i_d + 1, 0),
\]
We claim that
\[ P_{\pi^0 T, T}^\xi = P_{\alpha_0}^\xi T \cup P_{\alpha_1}^\xi T \cup \ldots \cup P_{\alpha_{d+1}}^\xi T. \]
which will prove that the partition \( P_{\pi^0 T, T}^\xi \) is finer than \( P_{d}^\xi T \).

Evidently, for all \( \omega \in \bigcup_{j=0}^{d+1} P_{\alpha_j}^\xi T \) condition (15) holds true, that is \( \omega \in P_{\pi^0 T, T}^\xi \).

Conversely, let \( \omega \in P_{\pi^0 T, T}^\xi \). If \( \xi(\omega) > \xi \circ T^{i_0+1}(\omega) \), then \( \omega \in P_{\alpha_0}^\xi T \).
Otherwise, let \( \tau = \max\{b \in \{0, \ldots, d\} \mid \xi \circ T^{i_b+1}(\omega) \geq \xi(\omega)\} \). Then \( \omega \in P_{\alpha_\tau}^\xi T \). This completes the proof of (13). \( \square \)

3.3. **Proof of Theorem 2.5.** Let \((\Omega, \mathcal{F}, \mu)\) be a probability space, \(T : \Omega \to \Omega\) be a measurable \(\mu\)-invariant transformation, and \(\Theta : \Omega \to \mathbb{R}^n\) be a measurable map such that \(\sigma(\{\Theta \circ T^k\}_{k \geq 0}) = \mathcal{F}\). We have to prove that
\[ h_{KS}^\mu(T) = \lim_{d \to \infty} \lim_{k \to \infty} \frac{1}{k} H_\mu(\tau_k(P_d^{\Theta, T})) \]
if either
(a) \(T\) is ergodic, or
(b) \(T\) is not ergodic, however \(\Omega\) can be embedded into some compact metrizable space so that \(\mathcal{F} = \mathcal{B}(\Omega)\).

In the case (a) it follows from Corollary 3.5 and the assumptions above that
\[ \Sigma_{\Theta, T}^{\circ} = \mathcal{F}, \]
which by Lemma 2.2 implies (17).

In the case (b) the equality (17) follows from the Ergodic decomposition theorem by the arguments of the proof of [10, Theorem 2.1]. \( \square \)

4. Residuality and Prevalence

4.1. **The set of non-injectivity.** Let \(\Theta : \Omega \to X\) be a continuous map between topological spaces \(\Omega\) and \(X\). Suppose also that \(\mu\) is a measure on the \(\sigma\)-algebra \(\mathcal{B}(\Omega)\) of Borel sets of \(\Omega\). In this section we give sufficient conditions on \(\Theta\) for the equivalence \(\sigma(\Theta) = \mathcal{B}(\Omega)\) and also prove Theorem 2.13.

The subset
\[ N_\Theta = \{ \omega \in \Omega \mid \Theta^{-1}\Theta(\omega) \neq \{\omega\} \} \]
of \(\Omega\) is called the *set of non-injectivity* of \(\Theta\). It plays a principal role in the further considerations.
We will now present a more useful description of $N_\Theta$. Given a set $\Omega$, a number $s \geq 1$, and map $\Theta : \Omega \to X$ into some set $X$ put

$$\Omega^s = \underbrace{\Omega \times \cdots \times \Omega}_s,$$

$$\Delta \Omega^s = \{(\omega, \ldots, \omega) \in \Omega^s | \omega \in \Omega\},$$

$$\Theta^s = \underbrace{\Theta \times \cdots \times \Theta}_s : \Omega^s \to X^s,$$

$$\Theta^s(\omega_1, \ldots, \omega_s) = (\Theta(\omega_1), \ldots, \Theta(\omega_s)).$$

In particular, for $s = 2$ consider the following subset of $\Omega^2$:

\begin{equation}
M = (\Theta^2)^{-1}(\Delta X^2) \setminus \Delta \Omega^2 = \{ (\omega, \omega') | \omega \neq \omega', \Theta(\omega) = \Theta(\omega') \}.
\end{equation}

Let also $p : \Omega^2 \to \Omega$ be the projection to the first coordinate. Then it is evident that

\begin{equation}
N_\Theta = p(M).
\end{equation}

The following lemma describes some properties of the set of non-injectivity.

**Lemma 4.1.** 1) $\Theta^{-1}(N_\Theta) = N_\Theta$.

2) For any subset $F \subset \Omega \setminus N_\Theta$ the restriction $\Theta|_F : F \to X$ is injective.

3) Suppose $\Omega$ and $X$ are Hausdorff, $\Omega$ is also second countable and locally compact (e.g. a manifold). Then $N_\Theta$ is an $F_\sigma$ subset of $\Omega$, and in particular $N_\Theta \in \mathcal{B}(\Omega)$.

**Proof.** Statements 1) and 2) are evident. Let us prove 3). It is easy to see that a topological space $X$ is Hausdorff iff the diagonal $\Delta X^2$ is closed in $X^2$. This implies that $(\Theta^2)^{-1}(\Delta X^2)$ is closed in $\Omega^2$, hence $M$ defined by (19) is second countable and locally compact as well.

Therefore $M = \bigcup_{i=1}^{\infty} M_i$ where each $M_i$ is compact. Hence

$$N_\Theta = p(M) = p\left( \bigcup_{i=1}^{\infty} M_i \right) = \bigcup_{i=1}^{\infty} p(M_i).$$

But each $p(M_i)$ is compact and so closed in $\Omega$. Hence $N_\Theta$ is an $F_\sigma$ set. \qed

Recall that a Polish space is a second countable completely metrizable topological space.

**Theorem 4.2.** Let $\Omega$ and $X$ be Polish spaces, $\mu$ be a measure on $\mathcal{B}(\Omega)$, $\Theta : \Omega \to X$ be a continuous map, and $N_\Theta$ be the set of its non-injectivity. Suppose $N_\Theta \in \mathcal{B}(\Omega)$ and $\mu(N_\Theta) = 0$. Then $\sigma(\Theta) \supseteq \mathcal{B}(\Omega)$. 

Proof. Since $\sigma(\Theta) \subset B(\Omega)$, it remains to consider the inverse inclusion. It suffices to show that for any open set $G \in B(\Omega)$ there exists some $\tilde{G} \in \sigma(\Theta)$ such that $\mu(G \triangle \tilde{G}) = 0$. Given $G$, put

$$\tilde{G} = G \setminus N_\Theta.$$ 

Then by 2) of Lemma 4.1 the restriction $\Theta|_{\tilde{G}} : \tilde{G} \to \Theta(\tilde{G})$ is one-to-one. So $\Theta(\tilde{G})$ is a one-to-one image of the Polish space $\tilde{G}$ under the continuous map $\Theta|_{\tilde{G}} : \tilde{G} \to X$. This implies, [14, Theorem 15.1], that $\Theta(\tilde{G}) \in B(X)$, whence

$$\tilde{G} = \Theta^{-1}(\Theta(\tilde{G})) \in \sigma(\Theta).$$

Therefore $\mu(G \triangle \tilde{G}) = \mu(G \cap N_\Theta) \leq \mu(N_\Theta) = 0$. □

4.2. Multijets transversality theorem. The proof of Theorem 2.13 is based on the so-called multijets transversality theorem, see [6, Chapter 2, Theorem 4.13]. We will formulate it below preserving the notation from [6].

Let $\Omega$ and $X$ be smooth manifolds, $\dim \Omega = m$, $\dim X = n$, $J^k(\Omega, X)$ be the manifold of $k$-jets of smooth maps $\Theta : \Omega \to X$,

$$\alpha : J^k(\Omega, X) \to \Omega$$

be the natural projection to the source,

$$\alpha^s = \alpha \times \cdots \times \alpha : J^k(\Omega, X)^s \to \Omega^s,$$

$$\Omega^s = \{(\omega_1, \ldots, \omega_s) \in \Omega^s \mid \omega_i \neq \omega_j \text{ for } i \neq j\},$$

and

$$J^k_s(\Omega, X) = (\alpha^s)^{-1}(\Omega^s).$$

Then $J^k_s(\Omega, X)$ is an open submanifold of $J^k(\Omega, X)^s$ and we have the map

$$j^k_s \Theta : \Omega^s \to J^k_s(\Omega, X), \quad j^k_s \Theta(\omega_1, \ldots, \omega_s) = (j^k \Theta(\omega_1), \ldots, j^k \Theta(\omega_s)).$$

The following result is called multijets transversality theorem.

**Theorem 4.3.** [6, Chapter 2, Theorem 4.13]. Let $B$ be a submanifold in $J^k_s(\Omega, X)$. Endow $C^\infty(\Omega, X)$ with the strong topology $C^\infty_S$. Then the set

$$\mathcal{V}_B = \{ \Theta \in C^\infty(\Omega, X) \mid j^k_s \Theta \text{ is transversal to } B \}$$

is residual in $C^\infty(\Omega, X)$. If $B$ is compact, then $\mathcal{V}_B$ is also open. □
We will apply this theorem to the case \( k = 0 \) and \( s = 2 \). First we will show that description (20) of \( N_\Theta \) is related to multijets transversality theorem. Recall that \( J^0(\Omega, X) = \Omega \times X \). Then

\[
J^0(\Omega, X) = \{(\omega, x, \omega', x') \mid \omega \neq \omega'\} \subset J^0(\Omega, X)^2 = (\Omega \times X)^2.
\]

Let \( \beta : J^0(\Omega, X) \rightarrow X^2 \) be the projection to the destination given by

\[
\beta(\omega, x, \omega', x') = (x, x'),
\]

and

\[
B = \beta^{-1}(\Delta X^2) = \{(\omega, x, \omega', x) \mid \omega \neq \omega'\}.
\]

Evidently, \( \beta \) is a submersion. Therefore it is transversal to \( \Delta X^2 \), and so \( B \) is a submanifold in \( J^0(\Omega, X) \) of codimension

\[
\text{codim} \Delta X^2 = \dim X^2 - \dim \Delta X^2 = \dim X = n.
\]

Also notice that \( B \) is non-compact. Consider the map

\[
j^0_2 \Theta : \Omega^{(2)} \rightarrow J^0(\Omega, X), \quad j^0_2 \Theta(\omega, \omega') = (\omega, \Theta(\omega), \omega', \Theta(\omega')).
\]

Then

\[
M = (j^0_2 \Theta)^{-1}(B) = \{ (\omega, \omega') \in \Omega^2 \mid \omega \neq \omega', \Theta(\omega) = \Theta(\omega') \}, \quad N_\Theta = p(M),
\]

as in (19) and (20).

**Corollary 4.4.** Let \( B \) be the submanifold of \( J^0(\Omega, X) \) given by (21) and

\[
\mathcal{V}_B = \{ \Theta \in C^\infty(\Omega, X) \mid j^0_2 \Theta \text{ is transversal to } B \}.
\]

Then by Theorem 4.3 \( \mathcal{V}_B \) is residual in \( C^\infty(\Omega, X) \) with respect to the strong topology \( C^\infty_S \). If \( m < n \), then for each \( \Theta \in \mathcal{V}_B \) the set \( N_\Theta \) has measure zero in the sense of Definition 2.11.

**Proof.** Notice that \( B \) has codimension \( n \) in \( J^0(\Omega, X) \). Therefore the submanifold \( M = (j^0_2 \Theta)^{-1}(B) \) has the same codimension \( n \) in \( \Omega^{(2)} \).

Since \( m < n \), we obtain that

\[
\dim M = \dim \Omega^{(2)} - n = 2m - n < m = \dim \Omega.
\]

Consider the restriction \( p|_M : M \rightarrow \Omega \). As \( \dim M < \dim \Omega \), each point of \( M \) is critical for \( p|_M \), whence by Sard’s theorem, [6], the image of the set of critical points of \( p|_M \), i.e. the set \( p(M) = N_\Theta \), has measure zero in the sense of Definition 2.11. □

**Corollary 4.5.** Suppose in Corollary 4.4 \( X = \mathbb{R}^n \) for some \( n \), so \( C^\infty(\Omega, \mathbb{R}^n) \) is a linear space. Then the set \( \mathcal{V}_B \) has a probe. In particular, by Lemmas 2.9 and 2.10 it is prevalent with respect to any of weak topologies \( C^r_W \) on \( C^\infty(\Omega, \mathbb{R}^n) \).
Proof. First we introduce some notation and prove Lemma 4.6 below. Let $M(n, k)$ be the space of $(n \times k)$-matrices ($n$ rows and $k$ columns) which can be identified with $\mathbb{R}^{nk}$, and $D_r(n, k)$ be the subset of $M(n, k)$ consisting of matrices of rank $r$. Then $D_r(n, k)$ is a smooth submanifold of codimension $(n - r)(k - r)$, see e.g. [15, Lemma 1.19]. Define the map

$$\Phi : M(n, k) \times M(n, k) \to M(2n, k),$$

which associate to each pair $A, B \in M(n, k)$ the matrix $\Phi(A, B)$ obtained by appending all rows of $B$ to $A$, see Figure 4.2. Evidently, $\Phi$ is a smooth diffeomorphism.

Now we can construct the probe for $\mathcal{V}_B$. Let $G : \Omega \to M(n, k)$ be a map satisfying statement of Lemma 4.6. For each $v \in \mathbb{R}^k$ define the following smooth map

$$L_v : \Omega \to \mathbb{R}^n, \quad L_v(\omega) = G(\omega)v,$$

and let

$$P = \{L_v \mid v \in \mathbb{R}^k\} \subset C^\infty(\Omega, \mathbb{R}^n).$$

Then $P$ is a linear subspace of $C^\infty(\Omega, \mathbb{R}^n)$ of dimension $\leq k$. We claim that $P$ is a probe for $\mathcal{V}_B$.

Indeed, let $\Theta \in C^\infty(\Omega, \mathbb{R}^n)$ be any map. For each $v \in \mathbb{R}^k$ denote

$$\Theta_v = \Theta + L_v,$$

so the translation of $P$ by $\Theta$ is the following affine subspace of $C^\infty(\Omega, \mathbb{R}^n)$:

$$\Theta + P = \{\Theta_v \mid v \in \mathbb{R}^k\}.$$ We should prove that the following set:

$$Q = \{v \in \mathbb{R}^k \mid j_2^0 \Theta_v \text{ is not transversal to } B\}$$

has Lebesgue measure zero in $\mathbb{R}^k$. Define the map

$$\Psi : \Omega^{(2)} \times \mathbb{R}^k \to J_0^2(\Omega, \mathbb{R}^n) \subset (\Omega \times \mathbb{R}^n)^2$$

by

$$\Psi(\omega, \omega', v) := j_2^0(\Theta + L_v)(\omega, \omega') = (\omega, \Theta(\omega) + L_v(\omega), \omega', \Theta(\omega') + L_v(\omega')).$$

Then the Jacobi matrix of $\Psi$ at point $(\omega, \omega', v)$ has the form shown in Figure 4.2, and so its rank is maximal and equals $2(m + n)$ due to the
choice of $G$. Hence $\Psi$ is a submersion. Therefore it is transversal to $B$, and so $\tilde{M} = \Psi^{-1}(B)$ is a submanifold in $\Omega^{(2)} \times \mathbb{R}^k$. Let $\pi : \tilde{M} \to \mathbb{R}^k$ be the restriction to $\tilde{M}$ of the natural projection $\Omega^{(2)} \times \mathbb{R}^k \to \mathbb{R}^k$. Then it is easy to see that $Q$ coincides with the set of critical values of $\pi$. Since $\pi$ is smooth, we get from Sard’s theorem that $Q$ has Lebesgue measure zero, see e.g. [7, Chapter 2, §3]. Corollary 4.5 is completed. □

Lemma 4.6. If $k \geq 2(m + n)$, then there exists a smooth map $G : \Omega \to M(n, k)$ such that for any pair of distinct points $\omega \neq \omega' \in \Omega$ the matrix

$$\Phi(G(\omega), G(\omega'))$$

has rank $2n$.

Proof. The proof is also based on the multijets transversality theorem. Consider the following spaces:

$$J^0(\Omega, M(n, k)) = \Omega \times M(n, k),$$
$$J^0_2(\Omega, M(n, k)) = \{(\omega, A, \omega', A') \mid \omega \neq \omega'\} \subset J^0(\Omega, M(n, k))^2 = (\Omega \times M(n, k))^2.$$

Let

$$\gamma = \Phi \circ \beta : J^0_2(\Omega, M(n, k)) \xrightarrow{\beta} M(n, k)^2 \xrightarrow{\Phi} M(2n, k)$$

be the projection to the destination $\beta$ composed with the diffeomorphism $\Phi$:

$$\gamma(\omega, A, \omega', A') = \Phi(A, A'),$$

and

$$\tilde{D}_r(2n, k) = \gamma^{-1}(D_r(2n, k)) = \{(\omega, A, \omega', A') \mid \omega \neq \omega', \text{ rank } \Phi(A, A') = r\}.$$ 

for $r < 2n$. Since $\gamma$ is a submersion, it is transversal to $D_r(2n, k)$, and so $\tilde{D}_r(2n, k)$ is a submanifold in $J^0_2(\Omega, M(n, k))$ of codimension

$$\text{codim } \tilde{D}_r(2n, k) = \text{codim } D_r(2n, k) = (2n - r)(k - r).$$

Then by multijets transversality theorem the set

$$T_r = \{G \in C^\infty(\Omega, \mathbb{R}^n) \mid j^0_2 G \text{ is transversal to } \tilde{D}_r(2n, k)\}$$
is residual. Hence, so is the intersection

\[ T = \bigcap_{r=0}^{2n-1} T_r, \]

and, in particular, \( T \) is non-empty. We claim that any \( G \in T \) satisfies the statement of the lemma.

First notice that the assumption \( k \geq 2(m + n) \) is equivalent to the inequality: \( k - 2n + 1 > 2m \). Then for \( 0 \leq r \leq 2n - 1 \) we have that

\[ \text{codim} \tilde{D}_r(2n, k) \geq \text{codim} \tilde{D}_{2n-1}(2n, k) = k - 2n + 1 > 2m = \text{dim} \Omega^{(2)}, \]

and so transversality of \( j^0_2G \) to \( \tilde{D}_r(2n, k) \) means that \( j^0_2G(\Omega^{(2)}) \cap \tilde{D}_r(2n, k) = \emptyset \). Thus if a map \( G : \Omega \rightarrow \mathbb{R}^n \) belongs to \( T \), then

\[ j^0_2G(\Omega^{(2)}) \cap \tilde{D}_r(2n, k) = \emptyset, \quad r = 0, 1, \ldots, 2n - 1. \]

This means that \( \text{rank} \Phi(G(\omega), G(\omega')) = 2n \) for any \( \omega \neq \omega' \in \Omega \). \( \Box \)

**Proof of Theorem 2.13.** Let \( \Omega \) be a smooth manifold of dimension \( m \), \( \mu \) be a Lebesgue absolute continuous measure on \( \mathcal{B}(\Omega) \), \( T : \Omega \rightarrow \Omega \) be a measurable \( \mu \)-invariant transformation, and \( n > m \).

Let \( V = V_B \) be defined by (22). Then by Corollaries 4.4 and 4.5 \( V \) is residual with respect to the strong topology \( C^\infty_S \) and prevalent with respect to the weak topology \( C^\infty_W \).

We claim that (9) holds for each \( \Theta \in V \). Indeed, by 3) of Lemma 4.1 \( N_\Theta \) is a Borel subset of \( \Omega \). Also by Corollary 4.4 it has measure zero in the sense of Definition 2.11. Since \( \mu \) is Lebesgue absolute continuous, we see that \( \mu(N_\Theta) = 0 \), whence by Theorem 4.2 \( \sigma(\Theta) \circ T = \mathcal{B}(\Omega) \).

Furthermore, as \( \Omega \) is an \( m \)-dimensional manifold, it can be embedded in \((2m + 1)\)-cube being a compact metric space. Therefore by (b) of Theorem 2.5 we have that

\[ h^K_S(\mu) = \lim_{d \to \infty} \lim_{k \to \infty} \frac{1}{k} \mu \left( \tau_k(P^0_\Theta, T) \right). \]

This completes Theorem 2.13. \( \Box \)

**References**

[1] J.M. Amigó, *The equality of Kolmogorov-Sinai entropy and metric permutation entropy generalized*, Physica D, **241**, (2012), no. 7, 789–793. MR MR2897545

[2] José María Amigó, *Permutation complexity in dynamical systems*, Springer Series in Synergetics, Springer-Verlag, Berlin, 2010, Ordinal patterns, permutation entropy and all that. MR 2583155 (2011f:37002)

[3] J.M. Amigó, M.B. Kennel, and L. Kocarev, *The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical systems*, Physica D, **210** (2005), 77–95.
[4] Christoph Bandt, Gerhard Keller, and Bernd Pompe, *Entropy of interval maps via permutations*, Nonlinearity 15 (2002), no. 5, 1595–1602. MR 1925429 (2003h:37048)

[5] C. Bandt, B. Pompe, *Permutation entropy: A natural complexity measure for time series*, Phys. Rev. Lett. 88, (2002), 174102.

[6] M. Golubitsky and V. Guillemin, *Stable mappings and their singularities*, Springer-Verlag, New York, 1973, Graduate Texts in Mathematics, Vol. 14. MR 0341518 (49 #6269)

[7] V. Guillemin and A. Polak *Differential topology*, Prentice-Hall, Englewood Cliff, NJ, 1974

[8] M. Hirsch, *Differential topology*, Graduate Texts in Mathematics, No. 33. Springer-Verlag, New York-Heidelberg, 1976. x+221 pp. MR MR0448362 (56 #6669)

[9] B. R. Hunt, T. Sauer, J. A. Yourke, *Prevalence: a translation-invariant “almost every” on infinite-dimensional spaces*, Bull. Amer. Math. Soc. 27, (1992), no. 2, 217-238.

[10] Karsten Keller, *Permutations and the Kolmogorov-Sinai entropy*, Discrete Contin. Dyn. Syst. 32 (2012), no. 3, 891–900. MR 2851883 (2012k:37020)

[11] K. Keller, A. Unakafov, and V. Unakafova, *On the Relation of KS Entropy and Permutation Entropy*, Physica D, 241 (2012), 1477–1481.

[12] K. Keller, M. Sinn, *Kolmogorov-Sinai entropy from the ordinal viewpoint*, Physica D, 239 (2010), 997–1000.

[13] K. Keller and M. Sinn, *A standardized approach to the Kolmogorov-Sinai entropy*, Nonlinearity 22 (2009), no. 10, 2417–2422. MR 2539761 (2010m:37010)

[14] A. S. Kechris, *Classical descriptive set theory*. Graduate Texts in Mathematics, 156. Springer-Verlag, New York, 1995. 402 p.

[15] J. Milnor, *Differential topology*, Mimeographed notes. Princeton University, New Jersey 1958.

[16] F. Takens, *Detecting strange attractors in turbulence*, in: Dynamical Systems and Turbulence (eds. D. A. Rand, L. S. Young), Lecture Notes in Mathematics 898, Springer-Verlag, Berlin-New York 1981, 366–381.

[17] T. Sauer, J. Yorke, and M. Casdagli, *Embeddology*, J. Stat. Phys., 65 (1991), 579–616.

[18] Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982. MR 648108 (84e:28017)