Observational Modeling of the Kolmogorov-Sinai Entropy

Uosef Mohammadi

Abstract. In this paper, Kolmogorov-Sinai entropy is studied using mathematical modeling of an observer $\Theta$. The relative entropy of a sub-$\sigma$-algebra having finite atoms is defined and then the ergodic properties of relative semi-dynamical systems are investigated. Also, a relative version of Kolmogorov-Sinai theorem is given. Finally, it is proved that the relative entropy of a relative $\Theta$-measure preserving transformations with respect to a relative sub-$\sigma_\Theta$-algebra having finite atoms is affine.

1. Introduction

In 1948, Shannon introduced the concept of entropy to information theory. The Shannon entropy is taken to indicate the degree of uncertainty ascribed to a random variable. Examining a random phenomenon as a member of a $\sigma$-algebra, Kolmogorov introduced the concept of entropy to ergodic theory in 1958. Kolmogorov’s entropy was improved by Sinai in [11]. Kolmogorov-Sinai entropy measures the rate of the loss of information for the iteration of finite partitions in a measure preserving transformation. Entropy as a mathematical device plays an important role in physical systems. On the other hand, one of the main objects in physical phenomena is the “observer”. So, a method is needed to measure the entropy of a system from the point of view of an observer. A modeling for an observer of a set $X$ is a fuzzy set $\Theta : X \rightarrow [0, 1]$ [6]. In fact these kinds of fuzzy sets are called one dimensional observers. In this paper, the Kolmogorov-Sinai entropy is studied using mathematical modeling of an observer. Any mathematical model according to the

2010 Mathematics Subject Classification. 37A35, 37B40.

Key words and phrases. Relative entropy, Relative semi-dynamical system, $m_\Theta$-equivalence, $m_\Theta$-generator, $(\Theta_1, \Theta_2)$-isomorphism.

Received: 20 January 2017, Accepted: 23 March 2017.

101
view point of an observer Θ is called a relative model [6, 7]. The notion of a relative semi-dynamical system as a generalization of a fuzzy dynamical system has been defined in [6]. Also, the concept of the entropy of a relative semi-dynamical system has been introduced in [5, 10].

This article is an attempt to present a new relative approach to the Kolmogorov-Sinai entropy. The first step in the evolution of relative entropy theory is to define the relative entropy of a sub-σ-Θ-algebra having finite atoms. Moreover, the definition of the relative entropy of a Θ-measure preserving transformations is based on the relative entropy of a sub-σ-Θ-algebra with finite atoms. Finally, the relative entropy of a relative semi-dynamical system is introduced and its ergodic properties are investigated.

2. Basic Notions

This section is devoted to provide the prerequisites that are necessary for the next section. Let (X, β) denotes a σ-finite measure space, i.e. a set equipped with a σ-algebra β of subsets of X. Further, let p denotes a probability measure on (X, β). Then (X, β, p) is called a probability space. Let φ : X → X be a measure preserving the invertible transformation of the probability space (X, β, p). In particular φ(β) = β and p(φ−1(A)) = p(A) for all A ∈ β. Then (X, β, p, φ) is called a dynamical system. The entropy of the partition ξ = {A1, ..., An} of the probability space (X, β, p) is defined by

$$H(\xi, p) = -\sum_{i=1}^{n} p(A_i) \log p(A_i),$$

and the entropy of the dynamical system (X, β, p, φ) with respect to the finite partition ξ is given by

$$h(φ, ξ, p) = \lim_{n→∞} \frac{1}{n} H(\xi ∨ φ^{-1}(ξ) ∨ ... ∨ φ^{-n}(ξ), p),$$

where φ−1(ξ) = {φ−1A1, ..., φ−1An}. Then the Kolmogorov-Sinai entropy of the automorphism φ is defined by

$$h(φ, p) = \sup_{ξ} h(φ, ξ, p),$$

where the supremum is taken over all finite partitions. In the following, we recall some known concepts of the relative structures.

Let Θ be an observer on X. Then we say λ ⊆ Θ if λ(x) ⊆ Θ(x) for all x ∈ X. Moreover, if λ1, λ2 ⊆ Θ then λ1 ∨ λ2 and λ1 ∧ λ2 are subsets of Θ, and defined by

$$(λ_1 ∨ λ_2)(x) = \sup \{λ_1(x), λ_2(x)\},$$

and

$$(λ_1 ∧ λ_2)(x) = \inf \{λ_1(x), λ_2(x)\}.$$
and

$$(\lambda_1 \land \lambda_2)(x) = \inf \{\lambda_1(x), \lambda_2(x)\},$$

where $x \in X$.

**Definition 2.1.** A collection $F_\Theta$ of subsets of $\Theta$ is said to be a $\sigma_\Theta$-algebra in $\Theta$ if $F_\Theta$ satisfies the following conditions [11],

(i) $\Theta \in F_\Theta$,

(ii) if $\lambda \in F_\Theta$ then $\lambda' = \Theta - \lambda \in F_\Theta$. $\lambda'$ is called the complement of $\lambda$ with respect to $\Theta$,

(iii) if $\{\lambda_i\}_{i=1}^\infty$ is a sequence in $F_\Theta$ then $\vee_{i=1}^\infty \lambda_i = \sup \lambda_i \in F_\Theta$,

(iv) $\frac{\Theta}{2}$ doesn’t belong to $F_\Theta$.

If $P_1$ and $P_2$ are two $\sigma_\Theta$-algebras on $X$ then $P_1 \lor P_2$ is the smallest $\sigma_\Theta$-algebra that contains $P_1 \cup P_2$, denoted by $[P_1 \cup P_2]$.

**Definition 2.2.** A positive $\Theta$-measure $m_\Theta$ over $F_\Theta$ is a function $m_\Theta : F_\Theta \rightarrow I$ which is countably additive. This means that if $\{\lambda_i\}$ is a disjoint countable collection of members of $F_\Theta$, (i.e. $\lambda_i \subseteq \lambda'_j = \Theta - \lambda_j$ whenever $i \neq j$) then

$$m_\Theta(\bigvee_{i=1}^\infty \lambda_i) = \sum_{i=1}^\infty m_\Theta(\lambda_i).$$

The $\Theta$-measure $m_\Theta$ has the following properties [11],

(i) $m_\Theta(\emptyset) = 0$,

(ii) $m_\Theta(\lambda' \lor \lambda) = m_\Theta(\Theta)$ and $m_\Theta(\lambda') = m_\Theta(\Theta) - m_\Theta(\lambda)$ for all $\lambda \in F_\Theta$,

(iii) $m_\Theta(\lambda \lor \mu) + m_\Theta(\lambda \land \mu) = m_\Theta(\lambda) + m_\Theta(\mu)$ for each $\lambda, \mu \in F_\Theta$,

(iv) $m_\Theta$ is a nondecreasing function i.e. if $\lambda, \eta \in F_\Theta$ and $\lambda \subseteq \Theta$, then $m_\Theta(\lambda) \leq m_\Theta(\eta)$.

The triple $(X, F_\Theta, m_\Theta)$ is called a $\Theta-$ measure space and the elements of $F_\Theta$ are called relative measurable sets. The $\Theta-$ measure space $(X, F_\Theta, m_\Theta)$ is called a relative probability $\Theta-$measure space if $m_\Theta(\Theta) = 1$ [11].

3. $\Theta$-RELATIONS AND ATOMS

**Definition 3.1.** Let $(X, F_\Theta, m)$ be a $\Theta-$measure space. The elements $\mu, \lambda$ of $F_\Theta$ are called $m_\Theta$-disjoint if $m_\Theta(\lambda \land \mu) = 0$.

A $\Theta-$relation ‘$\equiv_{(mod \ m_\Theta)}$’ on $F_\Theta$ is defined as bellow

$$\lambda = \mu(\mod m_\Theta) \iff m_\Theta(\lambda) = m_\Theta(\mu) = m_\Theta(\lambda \land \mu), \quad \lambda, \mu \in F_\Theta.$$

The $\Theta-$relation ‘$\equiv_{(mod \ m_\Theta)}$’ is an equivalence relation. $\tilde{F_\Theta}$ denotes the set of all equivalence classes induced by this relation, and $\tilde{\mu}$ is the
equivalence class determined by \( \mu \). For \( \lambda, \mu \in F_\Theta \), \( \lambda \wedge \mu = 0 \pmod{m_\Theta} \) iff \( \lambda, \mu \) are \( m_\Theta \)-disjoint. We shall identify \( \tilde{\mu} \) with \( \mu \).

**Definition 3.2.** Let \((X, F_\Theta, m_\Theta)\) be a \( \Theta \)--measure space, and \( P \) be a sub-\( \sigma_\Theta \)--algebra of \( F_\Theta \). Then an element \( \tilde{\lambda} \in \tilde{P} \) is an atom of \( P \) if

(i) \( m_\Theta(\lambda) > 0 \),

(ii) for each \( \tilde{\mu} \in \tilde{P} \) such that \( m_\Theta(\lambda \wedge \mu) = m_\Theta(\mu) \neq m_\Theta(\lambda) \) then \( m_\Theta(\mu) = 0 \).

**Theorem 3.3.** Let \((X, F_\Theta, m)\) be a \( \Theta \)--measure space, and \( P \) be a sub-
\( \sigma_\Theta \)--algebra of \( F_\Theta \). If \( \tilde{\lambda}_1, \tilde{\lambda}_2 \) are disjoint atoms of \( P \) then they are \( m_\Theta \)-disjoint.

**Proof.** Since \( \lambda_1 \wedge \lambda_2 \subseteq \lambda_1, \lambda_1 \wedge \lambda_2 \subseteq \lambda_2 \), and \( \lambda_1 \neq \lambda_2(\mod m_\Theta) \), we get \( \lambda_1 \wedge \lambda_2 \neq \lambda_i(\mod m_\Theta) \) for at least one \( i = 1, 2 \). Suppose \( \lambda_1 \wedge \lambda_2 \neq \lambda_2(\mod m_\Theta) \). Because \( \lambda_2 \) is an atom, \( \lambda_1 \wedge \lambda_2 = 0(\mod m_\Theta) \).

Now, we introduce \( R_\ast(F_\Theta) \) as bellow,

\( R_\ast(F_\Theta) = \{P : P \) is a sub -- \( \sigma_\Theta \) -- algebra of \( F_\Theta \) with finite atoms\}.

Assume that \( F_\Theta \) is a \( \sigma_\Theta \)--algebra, \( P_1, P_2 \in R_\ast(F_\Theta) \), and \( \{\lambda_i; i = 1, 2, \ldots, n\} \) and \( \{\mu_j; j = 1, \ldots, m\} \) denote the atoms of \( P_1 \) and \( P_2 \), respectively, then the atoms of \( P_1 \vee P_2 \) are \( \lambda_i \wedge \mu_j \) which \( m_\Theta(\lambda_i \wedge \mu_j) > 0 \) for each \( 1 \leq i \leq n, 1 \leq j \leq m \).

**Definition 3.4.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \( \Theta \)--measure space and \( P_1, P_2 \in R_\ast(F_\Theta) \). We say that \( P_2 \) is an \( m_\Theta \)--refinement of \( P_1 \), denoted by \( P_1 \leq m_\Theta P_2 \), if for each \( \mu \in P_2 \) there exists \( \lambda \in P_1 \) such that,

\[ m_\Theta(\lambda \wedge \mu) = m_\Theta(\mu) \]

**Theorem 3.5.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \( \Theta \)--measure space and \( P_1, P_2, P_3 \in R_\ast(F_\Theta) \). If \( P_1 \leq m_\Theta P_2 \) then,

\[ P_1 \vee P_3 \leq m_\Theta P_2 \vee P_3 \]

**Proof.** Let \( \mu \in \overline{P_2 \vee P_3} \). Then \( \mu = \lambda \wedge \gamma \) for some \( \lambda \in P_2 \) and \( \gamma \in P_3 \). Since \( P_1 \leq m_\Theta P_2 \), there exists \( \eta \in \tilde{P}_1 \) such that \( m_\Theta(\eta \wedge \lambda) = m_\Theta(\lambda) \).

Now,

\[ m_\Theta(\mu) = m_\Theta(\eta \wedge \gamma \wedge \lambda) \]
\[ = m_\Theta(\eta \wedge \lambda) + m_\Theta(\gamma) - m_\Theta((\eta \wedge \lambda) \vee \gamma) \]
\[ = m_\Theta(\lambda) + m_\Theta(\gamma) - m_\Theta((\eta \vee \lambda) \wedge (\lambda \vee \gamma)) \]
\[ = m_\Theta(\lambda) + m_\Theta(\gamma) - m_\Theta(\eta \vee \lambda) - m_\Theta(\lambda \vee \gamma) + m_\Theta(\eta \vee \gamma \vee \lambda) \]
\[ = m_\Theta(\lambda \wedge \gamma) \]
\[ = m_\Theta(\mu) \]
Hence, 
\[ m_\Theta(\eta \wedge \gamma \wedge \mu) = m_\Theta(\mu), \]
and the result follows. \( \square \)

4. Relative Entropy of a sub-\( \sigma_\Theta \)-algebra with Finite Atoms

**Definition 4.1.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \( \Theta \)-measure space, and \( P \) be a sub \( \sigma_\Theta \)-algebra of \( F_\Theta \) which \( P \in R_*(F_\Theta) \). The relative entropy of \( P \) is defined as

\[ H_\Theta(P, m_\Theta) = - \sum_{\mu \in P} m_\Theta(\mu) \log m_\Theta(\mu). \]

**Example 4.2.** Let \((X, \beta, p)\) be a classical probability measure space and \( \Theta = x_X \). Then \( F_\Theta = \{x_A : A \in \beta\} \) is a \( \sigma_\Theta \)-algebra on \( X \). Define \( m_\Theta(x_A) = p(A) \), \( A \in \beta \). Then \((X, F_\Theta, m_\Theta)\) is a relative probability \( \Theta \)-measure space. Let \( \alpha \) be a finite sub-\( \sigma \)-algebra of \( \beta \) and \( G = \{x_A : A \in \alpha\} \). So, \( G \in R_*(F_\Theta) \) and the relative entropy of \( G \) is given by

\[ H_\Theta(G, m_\Theta) = - \sum_{A \in G} m_\Theta(x_A) \log m_\Theta(x_A) \]

\[ = - \sum_{A \in G} p(A) \log p(A), \]

which is the Kolmogorov-Sinai entropy of the finite classical measurable sub-\( \sigma \)-algebra \( \alpha \) of the space \((X, \beta, p)\).

Thus, the concept of the relative entropy of a \( \sigma_\Theta \)-algebra with finite atoms is a generalization of the Kolmogorov-Sinai entropy of a finite measurable \( \sigma \)-algebra.

**Theorem 4.3.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \( \Theta \)-measure space, and \( P_1, P_2 \in R_*(F_\Theta) \) which \( P_1 = \{\lambda_i; 1 \leq i \leq n\} \) and \( P_2 = \{\mu_j; 1 \leq j \leq m\} \). If \( P_1 \approx_{m_\Theta} P_2 \) and \( P_1 \leq_{m_\Theta} P_2 \), then,

\[ H_\Theta(P_1, m_\Theta) \leq H_\Theta(P_2, m_\Theta). \]

**Proof.** Suppose that \( P_1 \leq_{m_\Theta} P_2 \). Then for each \( \mu_j \in P_2 \) there exists \( \lambda_k \in P_1 \) such that, \( m_\Theta(\mu_j \wedge \lambda_k) = m_\Theta(\mu_j) \). Since \( P_1 \approx_{m_\Theta} P_2 \) and \( \lambda_i \)'s are pairwise \( m_\Theta \)-disjoint, then

\[ m_\Theta(\mu_j) = m_\Theta(\mu_j \wedge (\lor \lambda_i)) \]

\[ = \sum_i m_\Theta(\mu_j \wedge \lambda_i). \]

Therefore, \( m_\Theta(\mu_j \wedge \lambda_i) = 0 \) for each \( i \neq k \). Hence,

\[ m_\Theta(\mu_j) \log m_\Theta(\mu_j) = \sum_i m_\Theta(\mu_j \wedge \lambda_i) \log m_\Theta(\mu_j \wedge \lambda_i). \]
Let \( P \)

Assume that

Then

\[
H_\Theta(P_2, m_\Theta) = -\sum_j \left( \sum_i m_\Theta(\lambda_i \land \mu_j) \log m_\Theta(\lambda_i \land \mu_j) \right)
\]

\[
= -\sum_{(i,j) \in \alpha} m_\Theta(\lambda_i \land \mu_j) \log m_\Theta(\lambda_i \land \mu_j)
\]

\[
\geq -\sum_{(i,j) \in \alpha} m_\Theta(\lambda_i \land \mu_j) \log m_\Theta(\lambda_i)
\]

\[
\geq -\sum_{i \in \beta} \log m_\Theta(\lambda_i) \sum_j m_\Theta(\lambda_i \land \mu_j).
\]

Since \( \mu_j \)'s are pairwise \( m_\Theta \)-disjoint, we have

\[
H_\Theta(P_2, m_\Theta) \geq -\sum_{i \in \beta} \log m_\Theta(\lambda_i)m_\Theta(\forall_j(\lambda_i \land \mu_j))
\]

\[
= -\sum_i m_\Theta(\lambda_i) \log m_\Theta(\lambda_i)
\]

\[
= H_\Theta(P_1, m_\Theta).
\]

\[\Box\]

**Definition 4.4.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \(\Theta\)-measure space and \(P_1, P_2 \in R_*(F_\Theta)\). We say that \(P_1\) and \(P_2\) are \(m_\Theta\)-equivalent, denoted by \(P_1 \approx_{m_\Theta} P_2\), if the following axioms are satisfied:

(i) If \(\lambda \in P_1\) then \(m_\Theta(\lambda \land (\forall \{\mu; \mu \in P_2\})) = m_\Theta(\lambda)\).

(ii) If \(\mu \in P_2\) then \(m_\Theta(\mu \land (\forall \{\lambda; \lambda \in P_1\})) = m_\Theta(\mu)\).

**Theorem 4.5.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \(\Theta\)-measure space, and \(P_1, P_2 \in R_*(F_\Theta)\). If \(P_1 \approx_{m_\Theta} P_2\) then,

\[P_1 \approx_{m_\Theta} P_1 \lor P_2\]

**Proof.** Assume that \(\bar{P}_1 = \{\lambda_i; 1 \leq i \leq n\}\) and \(\bar{P}_2 = \{\mu_j; 1 \leq j \leq m\}\).

We know that,

\[\bar{P}_1 \lor \bar{P}_2 = \{\lambda_i \land \mu_j; \lambda_i \in \bar{P}_1, \mu_j \in \bar{P}_2, m_\Theta(\lambda_i \land \mu_j) > 0\}\]

If \(\alpha = \{(i, j); V_{ij} = \lambda_i \land \mu_j \in \bar{P}_1 \lor \bar{P}_2\}\) then \(\alpha = \bigcup_i \{(i, j); j \in \beta_i\}\) where \(\beta_i = \{j; m_\Theta(V_{ij}) > 0\}\) and \(1 \leq i \leq n\). Note that if \(j \notin \beta_i\) then \(m_\Theta(V_{ij}) = 0\) and we have

\[
\forall_{i, j \in N} V_{ij} = \forall_{i \in N} (\forall_{j \in \beta_i} V_{ij})
\]

\[
= \forall_{1 \leq i \leq n} (\lambda_i \land (\forall_{j \in \beta_i} \mu_j)).
\]

Since the collections \(\{\lambda_i; 1 \leq i \leq n\}\) and \(\{\mu_j; 1 \leq j \leq m\}\) are \(m_\Theta\)-disjoint, then we have

\[
m_\Theta(\lambda_k \land (\forall_{i,j} V_{ij})) = m_\Theta(\lambda_k \land (\forall_{i} \lambda_i \land (\forall_{j \in \beta_i} \mu_j))\)
\]
\[
\begin{align*}
&= m_\Theta(\lambda_k \land (\forall j \in \beta_k \mu_j)) \\
&= m_\Theta(\lambda_k \land (\forall j \in \beta_k \mu_j)) \\
&= m_\Theta(\forall j \in \beta_k (\lambda_k \land \mu_j)) \\
&= \sum_{j \in \beta_k} m_\Theta(\forall V_{kj}) \\
&= \sum_j m_\Theta(\forall V_{kj}) \\
&= m_\Theta(\lambda_k \land (\forall j \mu_j)) \\
&= m_\Theta(\lambda_k).
\end{align*}
\]

**Theorem 4.6.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \(\Theta\)-measure space, and \(P_1, P_2 \in R_*(F_\Theta)\). If \(P_1 \approx_{m_\Theta} P_2\) then,

\[H_\Theta(P_1, m_\Theta) \leq H_\Theta(P_1 \lor P_2, m_\Theta).\]

**Proof.** Suppose that \(P_1 \approx_{m_\Theta} P_2\). By Theorem 3.8 we have \(P_1 \approx_{m_\Theta} P_1 \lor P_2\). Now suppose that \(\delta \in P_1 \lor P_2\). Then \(\delta = \lambda_i \land \mu_j\) which \(\lambda_i \in P_1\) and \(\mu_j \in P_2\). So for \(\lambda_i \in P_1\), \(m_\Theta(\delta) = m_\Theta(\delta \land \lambda_i)\) and therefore we have \(P_1 \leq_{m_\Theta} P_1 \lor P_2\). Now use Theorem 4.3. \(\square\)

**Theorem 4.7.** Let \((X, F_\Theta, m_\Theta)\) be a relative probability \(\Theta\)-measure space, and \(P_1, P_2 \in R_*(F_\Theta)\). If \(P_1 \approx_{m_\Theta} P_2\) then,

\[H_\Theta(P_1 \lor P_2, m_\Theta) \leq H_\Theta(P_1, m_\Theta) + H_\Theta(P_2, m_\Theta).\]

**Proof.** Suppose that \(g : [0, 1] \to \mathbb{R}\) be the convex function \(g(x) = x \log x\). Assume that \(P_1 = \{\lambda_i; 1 \leq i \leq n\}\) and \(P_2 = \{\mu_j; 1 \leq j \leq m\}\). Take \(\alpha_j = m_\Theta(\mu_j), 1 \leq j \leq m\) and for a fixed \(i (1 \leq i \leq n)\) put

\[x_j = \frac{m_\Theta(\lambda_i \land \mu_j)}{m_\Theta(\mu_j)}.
\]

We have,

\[
\sum_{j=1}^m \alpha_j x_j = \sum_{j=1}^m m_\Theta(\lambda_i \land \mu_j) \\
= m_\Theta(\lambda_i \land (\forall_{j=1}^m \mu_j)) \\
= m_\Theta(\lambda_i).
\]

Put

\[\eta = \{(i, j): 1 \leq i \leq n, 1 \leq j \leq m, m_\Theta(\lambda_i \land \mu_j) > 0\},\]
and $\beta_i = \{ j; m_\Theta (\lambda_i \land \mu_j) > 0 \}$. Let $\alpha = \sum_j \alpha_j$, then we get
\[ m_\Theta (\lambda_i) g \left( \frac{m_\Theta (\lambda_i)}{\alpha} \right) \leq \sum_{j=1}^{m} m_\Theta (\mu_j) g \left( \frac{m_\Theta (\lambda_i \land \mu_j)}{m_\Theta (\mu_j)} \right) \]
\[ = \sum_{j \in \beta_i} m_\Theta (\lambda_i \land \mu_j) \log \frac{m_\Theta (\lambda_i \land \mu_j)}{m_\Theta (\mu_j)}, \]
or
\[ m_\Theta (\lambda_i) \log \left( \frac{m_\Theta (\lambda_i)}{\alpha} \right) \leq \sum_{j \in \beta_i} m_\Theta (\lambda_i \land \mu_j) \log \left( \frac{m_\Theta (\lambda_i \land \mu_j)}{m_\Theta (\mu_j)} \right). \]

Now,
\[ H_\Theta (P_1, m_\Theta) = - \sum_{i=1}^{n} m_\Theta (\lambda_i) \log m_\Theta (\lambda_i) \]
\[ \geq - \sum_{i} m_\Theta (\lambda_i) \log \alpha - \sum_{i} \sum_{j \in \beta_i} m_\Theta (\lambda_i \land \mu_j) \log m_\Theta (\lambda_i \land \mu_j) \]
\[ + \sum_{i} \sum_{j \in \beta_i} m_\Theta (\lambda_i \land \mu_j) \log m_\Theta (\mu_j) \]
\[ = - \sum_{(i,j) \in \eta} m_\Theta (\lambda_i \land \mu_j) \log m_\Theta (\lambda_i \land \mu_j) \]
\[ + \sum_{j} \log m_\Theta (\mu_j) \sum_{i} m_\Theta (\lambda_i \land \mu_j) - \log \alpha \sum_{i} m_\Theta (\lambda_i) \]
\[ \geq H_\Theta (P_1 \lor P_2, m_\Theta) + \sum_{j} m_\Theta (\mu_j \land (\lor \lambda_i)) \log m_\Theta (\mu_j) \]
\[ = H_\Theta (P_1 \lor P_2, m_\Theta) - H_\Theta (P_2, m_\Theta). \]

Thus, $H_\Theta (P_1, m_\Theta) \leq H_\Theta (P_1 \lor P_2, m_\Theta)$. \(\square\)

5. RELATIVE ENTROPY OF A $\Theta$–MEASURE PRESERVING TRANSFORMATIONS

**Definition 5.1.** Suppose $(X, F_\Theta, m_\Theta)$ be an $\Theta$–measure space and $\Theta$ be a constant observer on $X$. A transformation $\varphi : (X, F_\Theta, m_\Theta) \to (X, F_\Theta, n_\Theta)$, is said to be a $\Theta$–measure preserving if $m_\Theta (\varphi^{-1}(\mu)) = n_\Theta(\mu)$ for all $\mu \in F_\Theta$.

**Theorem 5.2.** Suppose that 
\[ \varphi : (X, F_\Theta, m_\Theta) \to (X, F_\Theta, n_\Theta) \]
be a $\Theta$–measure preserving transformations. Then for each $P \in R_+(F_\Theta)$ we have,
\[ H_\Theta (P, m_\Theta) = H_\Theta (\varphi^{-1}(P), m_\Theta). \]
Proof. Since $\varphi$ is a $\Theta$–measure preserving, we have
$$m_{\Theta}(\varphi^{-1}(\mu)) = n_{\Theta}(\mu),$$
then,
$$H_{\Theta}(\varphi^{-1}(P), m_{\Theta}) = - \sum_{\mu \in P} m_{\Theta}(\varphi^{-1}(\mu)) \log m_{\Theta}(\varphi^{-1}(\mu))$$
$$= - \sum_{\mu \in P} n_{\Theta}(\mu) \log n_{\Theta}(\mu)$$
$$= H_{\Theta}(P, m_{\Theta}).$$
\[\square\]

**Definition 5.3.** Suppose $\varphi : (X, F_{\Theta}, m_{\Theta}) \to (X, F_{\Theta}, n_{\Theta})$, be a $\Theta$–measure preserving transformation. If $P \in R_{\ast}(F_{\Theta})$, we define the relative entropy of $\varphi$ with respect to $P$ as
$$h_{\Theta}(\varphi, P, m_{\Theta}) = \lim_{n \to \infty} \frac{1}{n} H_{\Theta}(\varphi^{-i}(P), m_{\Theta}).$$

**Theorem 5.4.** $\lim_{n \to \infty} \frac{1}{n} H_{\Theta}(\varphi^{-i}(P), m_{\Theta})$ exists.

**Proof.** Let
$$a_n = H_{\Theta}(\varphi^{-i}(P), m_{\Theta}) \geq 0.$$ Using Theorem 4.7 and Theorem 5.2, we have
$$a_{n+k} = H_{\Theta}(\varphi^{-i}(P), m_{\Theta})$$
$$\leq H_{\Theta}(\varphi^{-i}(P), m_{\Theta}) + H_{\Theta}(\varphi^{-i}(P), m_{\Theta})$$
$$= a_n + a_k.$$ So, for each $n, k$ we have $a_{n+k} \leq a_n + a_k$. Now, by Theorem 4.9 in [12]
$$\lim_{n \to \infty} \frac{a_n}{n}$$ exists. \[\square\]

**Theorem 5.5.** Let $\varphi : (X, F_{\Theta}, m_{\Theta}) \to (X, F_{\Theta}, n_{\Theta})$, be a $\Theta$–measure preserving transformations and $P \in R_{\ast}(F_{\Theta})$. Then,
(i) $h_{\Theta}(\varphi, \varphi^{-1}(P)) = h_{\Theta}(\varphi, P)$,
(ii) $h_{\Theta}(\varphi, \varphi^{-i}(P)) = h_{\Theta}(\varphi, P)$ for every $r \geq 1$.

**Proof.** (i) It is obvious.
(ii) We have
$$h_{\Theta}(\varphi, \varphi^{-i}(P), m_{\Theta}) = \lim_{n \to \infty} \frac{1}{n} H_{\Theta}(\varphi^{-i}(P), m_{\Theta})$$
$$= \lim_{n \to \infty} \frac{1}{n} H_{\Theta}(\varphi^{-i}(P), m_{\Theta})$$
Let $\phi$ be a relative semi-dynamical system denoted by $(X, F, \mu)$. Suppose $(X, F, \mu)$ is a relative probability measure space and $\varphi$ is a $\Theta$-measure preserving transformations.

**Theorem 5.6.** Let $\varphi : (X, F, \mu) \to (X, F, \mu)$, be a $\Theta$-measure preserving transformations and $P_1, P_2 \in R_s(F, \mu)$. If $P_1 \approx_r P_2$ and $P_1 \leq_m P_2$ then $h_{\Theta}(\varphi, P_1, \mu) \leq h_{\Theta}(\varphi, P_2, \mu)$.

**Proof.** The result follows from Theorem 6.4.

6. Relative Entropy and $(\Theta_1, \Theta_2)$-isomorphic Dynamical Systems

**Definition 6.1.** A relative semi-dynamical system is denoted by $(X, F, \mu, \varphi)$ which $(X, F, \mu, \varphi)$ is a relative probability $\Theta$-measure space and $\varphi$ is a $\Theta$-measure preserving transformations.

**Definition 6.2.** Let $(X, F, \mu, \varphi)$ be a relative semi-dynamical system and $L \in R_s(F, \mu)$. Suppose $[L]_\Theta$ denotes the $m_\Theta$-equivalence class induced by $L$. Then the relative entropy $h_{\Theta}(\varphi, [L]_\Theta)$ of $\varphi$ on $L$ is defined as

$$h_{\Theta}(\varphi, [L]_\Theta, \mu) = \sup_{P \in [L]_\Theta} h_{\Theta}(\varphi, P, \mu).$$

**Definition 6.3.** Suppose $(X_1, F_\Theta_1, m_\Theta_1)$ be a $\Theta_1$-measure space and $(X_2, F_\Theta_2, n_\Theta_2)$ be a $\Theta_2$-measure space. A transformation $\varphi : (X_1, F_\Theta_1, m_\Theta_1) \to (X_2, F_\Theta_2, n_\Theta_2)$, is said to be a $(\Theta_1, \Theta_2)$-measure preserving if

(i) $\varphi^{-1}(\mu) \in F_\Theta_1$ for every $\mu \in F_\Theta_2$, where $\varphi^{-1}(\mu)(x) = \mu(\varphi(x))$, $\forall x \in X$,

(ii) $m_\Theta_1(\varphi^{-1}(\mu)) = n_\Theta_2(\mu)$ for all $\mu \in F_\Theta_2$.

**Theorem 6.4.** Suppose $\varphi : (X_1, F_\Theta_1, m_\Theta_1) \to (X_2, F_\Theta_2, n_\Theta_2)$, be a $(\Theta_1, \Theta_2)$-measure preserving transformations. Then for each $P \in R_s(F_\Theta_2)$ we have,

$$H_{\Theta_2}(P, m_\Theta_2) = H_{\Theta_1}(\varphi^{-1}(P), m_\Theta_1).$$

**Proof.** By Theorem 6.2, the proof is clear.

**Definition 6.5.** A relative semi-dynamical system $\phi_1 = (X_1, F_\Theta_1, m_\Theta_1)$ is a $(\Theta_1, \Theta_2)$-factor of the relative semi-dynamical system $\phi_2 = (X_2, F_\Theta_2, n_\Theta_2)$ if there exists an onto $(\Theta_1, \Theta_2)$-measure preserving transformations (called homomorphism) $\psi : \phi_2 \to \phi_1$ such that,

$$\psi \circ \varphi_2 = \varphi_1 \psi.$$
Theorem 6.6. Let \( \phi_1 = (X_1, F_{\Theta_1}, m_{\Theta_1}) \) be a \((\Theta_1, \Theta_2)\)-factor of the relative semi-dynamical system \( \phi_2 = (X_2, F_{\Theta_2}, n_{\Theta_2}) \). Then for each \( L \in R_s(F_{\Theta}) \),

\[
h_{\Theta_1} \left( \varphi_1, \left[ \psi^{-1}(L) \right]_{\Theta_1}, m_{\Theta_1} \right) \leq h_{\Theta_2} \left( \varphi_2, \left[ L \right]_{\Theta_2}, m_{\Theta_2} \right),
\]

Where \( \psi : \phi_2 \to \phi_1 \) is the corresponding homomorphism.

Proof. Suppose that \( P \in [L]_{\Theta_2} \). Then by Theorem 6.4,

\[
H_{\Theta_2}(P, m_{\Theta_2}) = H_{\Theta_1}(\psi^{-1}(P), m_{\Theta_1}).
\]

Now,

\[
h_{\Theta_2}(\varphi_2, P, m_{\Theta_2}) = \lim_{n \to \infty} \frac{1}{n} H_{\Theta_2} \left( \bigvee_{i=0}^{n-1} \varphi_2^{-i}(P), m_{\Theta_2} \right)
= \lim_{n \to \infty} \frac{1}{n} H_{\Theta_1} \left( \psi^{-1} \left( \bigvee_{i=0}^{n-1} \varphi_2^{-i}(P) \right), m_{\Theta_1} \right)
= \lim_{n \to \infty} \frac{1}{n} H_{\Theta_1} \left( \bigvee_{i=0}^{n-1} \psi^{-1} \varphi_2^{-i}(P), m_{\Theta_1} \right)
= \lim_{n \to \infty} \frac{1}{n} H_{\Theta_1} \left( \bigvee_{i=0}^{n-1} \psi^{-1} \varphi_1^{-i}(P), m_{\Theta_1} \right)
= h_{\Theta_1} \left( \varphi_1, \psi^{-1}(P), m_{\Theta_1} \right).
\]

As \( P \) ranges over an \( m_{\Theta_2} \)-equivalence class \([L]_{\Theta_2}\) in \( R_s(F_{\Theta_2}) \), \( \psi^{-1}(P) \) ranges over a subset of the \( m_{\Theta_1} \)-equivalence class \([\psi^{-1}(L)]_{\Theta_1}\) in \( R_s(F_{\Theta_1}) \).

Definition 6.7. Two relative semi-dynamical systems \( \phi_1 = (X_1, F_{\Theta_1}, m_{\Theta_1}) \) and \( \phi_2 = (X_2, F_{\Theta_2}, m_{\Theta_2}) \) are said to be \((\Theta_1, \Theta_2)\)-isomorphic if there exists an invertible relative \((\Theta_1, \Theta_2)\)-measure preserving transformations \( \psi : \phi_1 \to \phi_2 \) (i.e both \( \psi \) and \( \psi^{-1} \) are relative measure preserving transformations) such that,

\[
\psi \circ \varphi_1 = \varphi_2 \circ \psi.
\]

The mapping \( \psi \) is called \((\Theta_1, \Theta_2)\)-isomorphism.

Theorem 6.8. Let \( \phi_1 \) and \( \phi_2 \) be \((\Theta_1, \Theta_2)\)-isomorphic semi-dynamical systems. Then for each \( L \in R_s(F_{\Theta_2}) \),

\[
h_{\Theta_1} \left( \varphi_1, \left[ \psi^{-1}(L) \right]_{\Theta_1}, m_{\Theta_1} \right) = h_{\Theta_2} \left( \varphi_2, \left[ L \right]_{\Theta_2}, m_{\Theta_2} \right),
\]

which \( \psi : \phi_1 \to \phi_2 \) is the corresponding \((\Theta_1, \Theta_2)\)-isomorphism.

Proof. The result follows from Theorem 6.6.
7. Relative Entropy and \( m_\Theta \)-generators of Relative Semi-Dynamical Systems

**Definition 7.1.** The relative entropy of the relative semi-dynamical system \((X, F_\Theta, m_\Theta, \varphi)\) is the number \( h_\Theta (\varphi, m_\Theta) \) defined by,
\[
h_\Theta (\varphi, m_\Theta) = \sup_P h_\Theta (\varphi, P, m_\Theta),
\]
where the supremum is taken over all sub-\(\sigma\)-algebras of \( F_\Theta \) which \( P \in R_+(F_\Theta) \).

**Definition 7.2.** \( P \in R_+(F_\Theta) \) is said to be an \( m_\Theta \)-generator of the relative semi-dynamical system \((X, F_\Theta, m_\Theta, \varphi)\) if there exists an integer \( r > 0 \) such that,
\[
\varphi_i Q \leq m_\Theta \forall i = 0 \varphi^{-1} P,
\]
for each \( Q \in R_+(F_\Theta) \).

**Theorem 7.3.** If \( P \) is an \( m_\Theta \)-generator of the relative semi-dynamical system \((X, F_\Theta, m_\Theta, \varphi)\) then,
\[
h_\Theta (\varphi, Q, m_\Theta) \leq h_\Theta (\varphi, P, m_\Theta),
\]
for each \( Q \in R_+(F_\Theta) \).

**Proof.** Let \( Q \in R_+(F_\Theta) \) be any arbitrary sub-\(\sigma\)-algebra of \( F_\Theta \). Since \( P \) is an \( m_\Theta \)-generator, \( Q \leq m_\Theta \forall i = 0 \varphi^{-1} P \) follows from Theorem 7.1.
\[
h_\Theta (\varphi, Q, m_\Theta) \leq h_\Theta (\varphi, \varphi_i \forall i = 0 \varphi^{-1} P, m_\Theta) = h_\Theta (\varphi, P, m_\Theta).
\]
\[\square\]

Now we can deduce the following version of the Kolmogorov-Sinai theorem.

**Theorem 7.4.** If \( P \) is an \( m_\Theta \)-generator of the relative semi-dynamical system \((X, F_\Theta, m_\Theta, \varphi)\) then,
\[
h_\Theta (\varphi, m_\Theta) = h_\Theta (\varphi, P, m_\Theta).
\]

**Proof.** It is obvious. \[\square\]

**Theorem 7.5.** Let \((X, F_\Theta, m_\Theta, \varphi)\) be a relative semi-dynamical system. Then, the map \( m_\Theta \mapsto h_\Theta (\varphi, m_\Theta) \) is affine, i.e,
\[
h_\Theta (\varphi, \lambda m_\Theta + (1 - \lambda) n_\Theta) = \lambda h_\Theta (\varphi, m_\Theta) + (1 - \lambda) h_\Theta (\varphi, n_\Theta),
\]
for each pair \( m_\Theta \) and \( n_\Theta \) of the relative probability \( \Theta \)-measures and \( \lambda \in [0, 1] \).
Proof. Suppose that \( P \in R_\Theta(F_\Theta) \). If \( m_\Theta \) and \( n_\Theta \) are two relative probability \( \Theta \)-measures and \( \lambda \in [0,1] \) then,

\[
H_\Theta(P, \lambda m_\Theta + (1 - \lambda) n_\Theta) \geq \lambda H_\Theta(P, m_\Theta) + (1 - \lambda) H_\Theta(P, n_\Theta).
\]

The ‘concavity’ inequality (7.1) is a direct consequence of the definition of \( H_\Theta(P, m_\Theta) \) and the ‘concavity’ of the function \( x \mapsto -x \log x \).

Conversely, one has inequalities

\[
- \log (\lambda m_\Theta(\mu_i) + (1 - \lambda) n_\Theta(\mu_i)) \leq - \log \lambda - \log m_\Theta(\mu_i),
\]

and

\[
- \log (\lambda m_\Theta(\mu_i) + (1 - \lambda) n_\Theta(\mu_i)) \leq - \log (1 - \lambda) - \log n_\Theta(\mu_i).
\]

Because \( x \mapsto - \log x \) is decreasing, therefore, one obtains the ‘convexity’ bound,

\[
H_\Theta(P, \lambda m_\Theta + (1 - \lambda) n_\Theta) \leq \lambda H(P, m_\Theta) + (1 - \lambda) H(P, n_\Theta)
- \lambda \log \lambda - (1 - \lambda) \log (1 - \lambda).
\]

Now replacing \( P \) by \( \varphi^{n-1}_i \varphi^{-i}(P) \) in (7.1), dividing by \( n \) and taking the limit \( n \to \infty \) gives

\[
h_\Theta(\varphi, P, \lambda m_\Theta + (1 - \lambda) n_\Theta) \geq \lambda h_\Theta(\varphi, P, m_\Theta) + (1 - \lambda) h_\Theta(\varphi, P, n_\Theta).
\]

Similarly from (7.2), since

\[
- \frac{(\lambda \log \lambda + (1 - \lambda) \log (1 - \lambda))}{n} \to 0,
\]

as \( n \to \infty \), one deduces the converse inequality

\[
h_\Theta(\varphi, P, \lambda m_\Theta + (1 - \lambda) n_\Theta) \leq \lambda h_\Theta(\varphi, P, m_\Theta) + (1 - \lambda) h_\Theta(\varphi, P, n_\Theta).
\]

Hence one concludes that the map \( m_\Theta \mapsto h_\Theta(\varphi, P, m_\Theta) \) is affine. Finally, it follows from Theorem 7.4 that the relative entropy is affine. \( \square \)

This is somewhat surprising and is of great significance in the application of the relative entropy.

8. Concluding Remarks and Open Problems

In this paper, the notion of the relative entropy for a sub-\( \sigma_\Theta \)-algebra with finite atoms is presented. The entropy of a relative semi-dynamical system is defined using the observer notion and its properties are investigated. Also, the notion of an \( m_\Theta \)-generator for a relative semi-dynamical system is introduced and a relative version of Kolmogorov-Sinai theorem concerning the entropy of a relative semi-dynamical system is given. Finally, it is proved that the relative entropy of a relative \( \Theta \)-measure preserving transformations with respect to a relative sub-\( \sigma_\Theta \)-algebra having finite atoms is affine.
An interesting open problem is to establish a theorem on existence of \( m_{\Theta} \)-generators for relative semi-dynamical systems.

**References**


Department of Mathematics, Faculty of Science, University of Jiroft, Jiroft, Iran.

E-mail address: u.mohamadi@ujiroft.ac.ir