RELATIVE ENTROPY OF RELATIVE MEASURE PRESERVING MAPS WITH CONSTANT OBSERVERS

MOLAEI M.R. & GHAZANFARI B.*
DEPARTMENT OF MATHEMATICS
SHAHID BAHONAR UNIVERSITY OF KERMAN
P.O.BOX 76135-133, KERMAN, IRAN
E-MAIL: MRMOLAEI@MAIL.UK.AC.IR
* DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE, LORESTAN UNIVERSITY
P.O. BOX 465, KHORAMABAD, IRAN
E-MAIL: GHAZANFARI.B@LU.AC.IR

(Received: 12 June 2007)

Abstract. In this paper the notion of relative entropy for a finite relative partition is considered. Relative measure preserving map on a relative probability space with a constant observer is studied. The notion of relative entropy for relative measure preserving maps is considered. Relative probability spaces and relative entropy as a generalization of the notions of fuzzy probability space and the entropy of fuzzy dynamical systems are studied.

AMS Classification: 37A99, 28D99
Keywords: Relative entropy, Relative probability space, Relative measure preserving map

1. INTRODUCTION

In this paper we assume that X is a non-empty set and μ is an observer of X [3, 4, 6], i.e. μ is a function from X to $[0, 1]$.
In section 2 we consider basic notions of the relative probability spaces. This definition can be considered as a generalization of the one which has been presented in [2, 5]. We are going to study 'relative entropy' of a finite $\mu-$partition and relative entropy of a relative measure preserving maps in sections 3 and 4 respectively. These notions generalize the notions of entropy and conditional entropy of fuzzy dynamical systems [1]. In section 5 we will consider the relative entropy and generators of the relative measure preserving maps.

2. BASIC NOTIONS

Let μ be an observer on X. Then we say $\lambda \subseteq \eta$ if $\lambda(x) \leq \eta(x)$ for all $x \in X$. Moreover if $\lambda_1, \lambda_2 \subseteq \mu$ then $\lambda_1 \cup \lambda_2$ and $\lambda_1 \cap \lambda_2$ are subsets of μ, and defined by :

$$(\lambda_1 \cup \lambda_2)(x) = \sup\{\lambda_1(x), \lambda_2(x)\}$$

and

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

where $x \in X$. Product and difference of subsets μ are defined by

$$(\lambda\eta)(x) = \lambda(x) \cdot \eta(x),\ (\lambda - \eta)(x) = Max\{\lambda(x) - \eta(x), 0\},\ where\ x \in X.$$

Let us introduce relative measure space or briefly $\mu-$measure space.

Definition 2.1. *A collection $\mathcal{M}_\mu$ of subsets of μ is said to be a $\sigma_\mu-$algebra in μ if $\mathcal{M}_\mu$ has the following properties* :
(i) $\mu \in \mathcal{M}_\mu$,
(ii) If $\lambda \in \mathcal{M}_\mu$, then $\lambda' = \mu - \lambda \in \mathcal{M}_\mu$. λ' is the complement of λ with respect to μ or relative complement of λ.
(iii) If $\lambda_n \in \mathcal{M}_\mu$ for $n = 1, 2, ...$ then $\lambda = \cup_{n=1}^{\infty}\lambda_n \in \mathcal{M}_\mu$, where $\lambda(x) = \sup\{\lambda_n(x) : n = 1, 2, ...\}\ \forall x \in X$,
(iv) $\mu/2$ doesn't belong to $\mathcal{M}_\mu$.

The relation $\subseteq$ is a partial ordering on $\mathcal{M}_\mu$.
Relative complement relation satisfies the following conditions:
(i) $(\lambda')' = \lambda$ for every $\lambda \in \mathcal{M}_\mu$, *(ii)* If $\lambda \subset \eta$ then $\eta' \subset \lambda'$ for every $\lambda, \eta \in \mathcal{M}_\mu$.
So $\mathcal{M}_\mu$ is a $\sigma_\mu-$lattice.

Definition 2.2. *A positive $\mu-$measure is a function $m_\mu : \mathcal{M}_\mu \to [0, \infty)$ which is countably additive. This means that if $\{\lambda_n\}$ is a disjoint countable collection of members of $\mathcal{M}_\mu$, (i.e. $\lambda_i \subseteq \lambda'_j = \mu - \lambda_j$ whenever $i \neq j$) then*

$$m_\mu(\cup_{n=1}^{\infty}\lambda_n) = \sum_{n=1}^{\infty} m_\mu(\lambda_n).$$

The $\mu-$measure $m_\mu : \mathcal{M}_\mu \longrightarrow [0, \infty)$ has the following properties:
(i) $m_\mu(\chi_\emptyset) = 0$.
(ii) $m_\mu(\lambda) \geq 0$ for all $\lambda \in \mathcal{M}_\mu$.
(iii) $m_\mu(\lambda' \cup \lambda) = m_\mu(\mu)$, and $m_\mu(\lambda') = m_\mu(\mu) - m_\mu(\lambda)$ for all $\lambda \in \mathcal{M}_\mu$.
(iv) $m_\mu(\lambda \cap \eta) + m_\mu(\lambda \cup \eta) = m_\mu(\lambda) + m_\mu(\eta)$.
(v) m_μ is a nondecreasing function i.e. if $\lambda, \eta \in \mathcal{M}_\mu$ and $\lambda \subseteq \eta$,
then $m_\mu(\lambda) \leq m_\mu(\eta)$.
(vi) Let $\lambda \in \mathcal{M}_\mu$ be given. Then '$m_\mu(\lambda \cap \eta) = m_\mu(\eta)$ for all $\eta \in \mathcal{M}_\mu$' if and only if $m_\mu(\lambda) = m_\mu(\mu)$.
(vii) $m_\mu(\lambda \cap \eta) = 0$, if λ, and η are pairwise disjoint in $\mathcal{M}_\mu$.
(viii) m_μ is a $\sigma_\mu-$subadditive function, i.e.

$$m_\mu(\cup_{n=1}^{\infty}\lambda_n) \leq \sum_{n=1}^{\infty} m_\mu(\lambda_n)$$

for any sequence $\{\lambda_n\}_{n=1}^{\infty} \subset \mathcal{M}_\mu$.
We call the triple $(X, \mathcal{M}_\mu, m_\mu)$ a relative probability space if $m_\mu(\mu) = 1$.

Example 2.3. *Let $(X, \mathcal{B}, p)$ be a probability space in the crisp sense. Suppose $\mu : X \longrightarrow [0,1]$ is an observer, and $\mathcal{M}_\mu = \{\lambda_E = \chi_E \cdot \mu : E \in \mathcal{B}\}$, where χ_E is the characteristic function of the set $E \subset X$ and $m_\mu : \mathcal{M}_\mu \longrightarrow [0,1]$ defined by $m_\mu(\lambda_E) = p(E)$. Then $(X, \mathcal{M}_\mu, m_\mu)$ is a relative probability space.*

Example 2.4. *Let $X = [-1,1]$ and $\mu : X \longrightarrow [0,1]$ be defined by $\mu(x) = |x|$. Suppose $\lambda \subset \mu$ be defined by:*

$$\lambda(x) = \begin{cases} \frac{|x|}{4} & if\ x \in [-1,0) \\ \frac{3|x|}{4} & if\ x \in [0,1]. \end{cases}$$

Moreover let $\mathcal{M}_\mu = \{\chi_\emptyset, \mu, \lambda, \lambda', \lambda \cap \lambda', \lambda \cup \lambda'\}$. Then $m_\mu : \mathcal{M}_\mu \longrightarrow [0,1]$ defined by $m_\mu(\mu) = 1 = m_\mu(\lambda \cup \lambda')$, $m_\mu(\chi_\emptyset) = m_\mu(\lambda \cap \lambda') = 0$ and $m_\mu(\lambda) = m_\mu(\lambda') = 1/2$, is a relative probability measure.

A pairwise disjoint family $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$ of elements of $\mathcal{M}_\mu$ is called a $\mu-$measurable finite partition of μ if

$$m_\mu(\cup_{j=1}^n \lambda_j) = m_\mu(\mu).$$

We denote the set of all $\mu-$measurable partitions of μ by $\mathbf{P}$.

Let $P,\ Q \in \mathbf{P}$. We say that Q is a refinement of P, and we write $P \prec Q$ if every element of P is a union of some elements of Q.
If $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $Q = \{\eta_1, \eta_2, ..., \eta_m\}$, then $P \vee Q$ is the set $\{\lambda_i \cap \eta_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$.

Proposition 2.5. *If P and Q are finite $\mu-$measurable partitions of μ then $P \vee Q$ is also a finite $\mu-$measurable partition and $P \prec P \vee Q$.*

Proof. Let $P, Q \in \mathbf{P}$, $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$, and $Q = \{\eta_1, \eta_2, ..., \eta_m\}$.
We know that: $\lambda_i \subseteq \mu - \lambda_k,\ ,\eta_j \subseteq \mu - \eta_l,\ \forall\ k \neq i,\ and \ \forall\ l \neq j$. Hence

$$\lambda_i \cap \eta_j \subseteq (\mu - \lambda_k) \cap (\mu - \eta_l) = \lambda'_k \cap \eta'_l.$$

Since $(\lambda_k \cap \eta_l) \subseteq \lambda_k$, and

$$(\lambda_k \cap \eta_l) \subseteq \eta_l,\ then\ \lambda'_k \subseteq (\lambda_k \cap \eta_l)',\ \eta'_l \subseteq (\lambda_k \cap \eta_l)'.$$

So

$$\lambda'_k \cap \eta'_l \subseteq (\lambda_k \cap \eta_l)' = \mu - (\lambda_k \cap \eta_l).$$

Therefore, $\lambda_i \cap \eta_j \subseteq \mu - (\lambda_k \cap \eta_l),\ \forall\ k \neq i,\ \forall\ l \neq j$.
We also have:

$$\begin{aligned} m_\mu(\cup_{i=1}^n \cup_{j=1}^m (\lambda_i \cap \eta_j)) &= \sum_{i=1}^n m_\mu(\lambda_i \cap (\cup_{j=1}^m \eta_j)) = m_\mu((\cup_{i=1}^n \lambda_i) \cap (\cup_{j=1}^m \eta_j)) \\ &= \sum_{i=1}^n m_\mu(\lambda_i) = m_\mu(\cup_{i=1}^n \lambda_i) = m_\mu(\mu). \end{aligned}$$

The third equality follows from the following equalities.

$m_\mu((\cup_{i=1}^{n}\lambda_i)\cap(\cup_{j=1}^{m}\eta_j))+m_\mu((\cup_{i=1}^{n}\lambda_i)\cup(\cup_{j=1}^{m}\eta_j))=m_\mu(\cup_{i=1}^{n}\lambda_i)+m_\mu(\cup_{j=1}^{m}\eta_j)$.
So

$$m_\mu((\cup_{i=1}^{n}\lambda_i)\cap(\cup_{j=1}^{m}\eta_j))+m_\mu(\mu)=m_\mu(\mu)+m_\mu(\mu).$$

□

On can deduce that: fuzzy probability space [2, 5] is a relative probability space when $\mu=\chi_X$.

3. RELATIVE ENTROPY OF A FINITE $\mu-$PARTITION

We begin this section with the definition of $\mu-$measurable partition.
Let $(X,\mathcal{M}_\mu,m_\mu)$ be a relative probability space. We say the $\mu-$partition P of μ is a $\mu-$measurable partition if $P\subseteq\mathcal{M}_\mu$.
Let $P=\{\lambda_1,\lambda_2,...,\lambda_n\}$ be a $\mu-$measurable partition of μ. Then the relative entropy $H(P)$ of P is defined by Shannon's formula:

$$H(P)=-\sum_{i=1}^{n}S(m_\mu(\lambda_i))$$

where

$$S(x)=\begin{cases} x\log x & if\ x>0\\ 0 & if\ x=0\end{cases}$$

So

$$H(P)=-\sum_{i=1}^{n}m_\mu(\lambda_i)\log m_\mu(\lambda_i),$$

where $0\log 0$ is 0.

Example 3.1. *Let $(X,\mathcal{M}_\mu,m_\mu)$ be the relative probability space of Example 2.4. Then $P=\{\lambda,\lambda'\}$ is a $\mu-$measurable partition of μ. It has the relative entropy $H(P)=\log 2$.*

We define the relative conditional probability measure $m_\mu(\cdot\mid\eta):\mathcal{M}_\mu\longrightarrow[0,1]$ by $m_\mu(\lambda\mid\eta)=\frac{m_\mu(\lambda\cap\eta)}{m_\mu(\eta)}$ for all $\lambda,\eta\in\mathcal{M}_\mu$.

Theorem 3.2. *The relative entropy $H:\mathbf{P}\longrightarrow\mathbb{R}$ has the following properties:*
(i) $H(P)\geq 0$ for every $P\in\mathbf{P}$.
(ii) If $P,Q\in\mathbf{P}$, and $P\prec Q$, then $H(P)\leq H(Q)$.
(iii) $H(P)\leq H(P\vee Q)$ for every $P,Q\in\mathbf{P}$.

Proof. (i) is evident.
(ii) If $P=\{\lambda_1,\lambda_2,...,\lambda_n\},\ Q=\{\eta_1,\eta_2,...,\eta_m\},$ *and* $P\prec Q$, then for each $\eta_j\in Q$ there exists λ_{i_0} such that $\eta_j\subseteq\lambda_{i_0}$. Since P is a pairwise disjoint $\mu-$measurable partition, we have $\eta_j=\eta_j\cap\lambda_{i_0}\subseteq\mu-\lambda_i$ for every $i\in\{1,2,...,n\}$, with $i\neq i_0$.

If we denote, $\Gamma = \{(i,j) : m_\mu(\lambda_i \cap \eta_j) > 0,\ i = 1,...,n,\ j = 1,...,m\}$, and $\gamma = \{i : m_\mu(\lambda_i) > 0,\ i = 1,...,n\}$, then

$$\begin{aligned} H(Q) &= -\sum_{(i,j)\in\Gamma} m_\mu(\lambda_i \cap \eta_j) \log m_\mu(\lambda_i \cap \eta_j) \\ &= -\sum_{(i,j)\in\Gamma} m_\mu(\lambda_i \cap \eta_j) \log m_\mu(\lambda_i \mid \eta_j) - \sum_{(i,j)\in\Gamma} m_\mu(\lambda_i \cap \eta_j) \log m_\mu(\lambda_i) \\ &\geq -\sum_{i\in\gamma} \log m_\mu(\lambda_i) \sum_{j=1}^{m} m_\mu(\lambda_i \cap \eta_j) = -\sum_{i\in\gamma} \log m_\mu(\lambda_i) m_\mu(\lambda_i \cap (\cup_{j=1}^{m} \eta_j)) \\ &= -\sum_{i\in\gamma} m_\mu(\lambda_i) \log m_\mu(\lambda_i) = H(P). \end{aligned}$$

(iii) Since $P \prec P \vee Q$, then (ii) implies to (iii). □

Definition 3.3. *Let $\eta \in \mathcal{M}_\mu$ and $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$ be a $\mu-$measurable partition of μ. Then the relative conditional entropy of P given η is defined by:*

$$H(P \mid \eta) = -\sum_{i=1}^{n} m_\mu(\lambda_i \mid \eta) \log m_\mu(\lambda_i \mid \eta)$$

where $m_\mu(\lambda_i \mid \eta) = \frac{m_\mu(\lambda_i \cap \eta)}{m_\mu(\eta)}$ when $m_\mu(\eta) > 0$.

Definition 3.4. *If $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $Q = \{\eta_1, \eta_2, ..., \eta_m\}$ are two $\mu-$measurable partitions of μ and*

$$m_\mu^+(\eta_j \mid \lambda_i) = \begin{cases} m_\mu(\eta_j \mid \lambda_i) & if\ m_\mu(\lambda_i) > 0 \\ 0 & if\ m_\mu(\lambda_i) = 0 \end{cases}.$$

then we define

$$\begin{aligned} H(Q \mid P) &= -\sum_{i=1}^{n}\sum_{j=1}^{m} m_\mu(\lambda_i) S(m_\mu^+(\eta_j \mid \lambda_i)) \\ &= -\sum_{i=1}^{n}\sum_{j=1}^{m} m_\mu(\lambda_i) m_\mu(\eta_j \mid \lambda_i) \log m_\mu(\eta_j \mid \lambda_i). \end{aligned}$$

Theorem 3.5. *Let $P, Q, R \in \mathbf{P}$. Then the following properties hold:*
(i) $H(Q \mid P) \geq 0$.
(ii) $H(Q \vee R \mid P) = H(R \mid P \vee Q) + H(Q \mid P)$.
(iii) If $P \prec Q$ then $H(P \mid R) \leq H(Q \mid R)$, $\forall R \in \mathbf{P}$.
(iv) If $P \prec Q$ then $H(R \mid P) \geq H(R \mid Q)$, $\forall R \in \mathbf{P}$.
(v) $H(R \vee Q \mid P) \leq H(R \mid P) + H(Q \mid P)$.

Proof. If $P = \{\lambda_1, \lambda_2, ..., \lambda_n\}$, $Q = \{\eta_1, \eta_2, ..., \eta_m\}$ and $R = \{\gamma_1, \gamma_2, ..., \gamma_r\}$, then
(i) is obvious.
(ii) If $m_\mu(\lambda_i \cap \eta_j) > 0$, then

$$m_\mu(\eta_j \cap \gamma_k \mid \lambda_i) = \frac{m_\mu(\lambda_i \cap \eta_j \cap \gamma_k)}{m_\mu(\lambda_i \cap \eta_j)} \cdot \frac{m_\mu(\lambda_i \cap \eta_j)}{m_\mu(\lambda_i)} = m_\mu(\gamma_k \mid \lambda_i \cap \eta_j) \cdot m_\mu(\eta_j \mid \lambda_i).$$

So

$$\begin{aligned}
H(Q \vee R \mid P) &= -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} m_\mu(\lambda_i) S(m_\mu^+(\eta_j \cap \gamma_k \mid \lambda_i)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} m_\mu(\lambda_i) S(m_\mu^+(\gamma_k \mid \lambda_i \cap \eta_j) \cdot m_\mu^+(\eta_j \mid \lambda_i)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} m_\mu(\lambda_i) m_\mu^+(\eta_j \mid \lambda_i) S(m_\mu^+(\gamma_k \mid \eta_j \cap \lambda_i)) \\
&\quad -\sum_{i=1}^{n}\sum_{j=1}^{m} m_\mu(\lambda_i) \sum_{k=1}^{r} m_\mu^+(\gamma_k \mid \lambda_i \cap \eta_j) S(m_\mu^+(\eta_j \mid \lambda_i)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} m_\mu(\lambda_i \cap \eta_j) S(m_\mu^+(\gamma_k \mid \lambda_i \cap \eta_j)) \\
&\quad -\sum_{i=1}^{n}\sum_{j=1}^{m} m_\mu(\lambda_i) S(m_\mu^+(\eta_j \mid \lambda_i)) \\
&= H(R \mid P \vee Q) + H(Q \mid P).
\end{aligned}$$

(iii) The proposition (ii) implies

$$H(Q \mid R) = H(Q \vee P \mid R) = H(P \mid R) + H(Q \mid P \vee R) \geq H(P \mid R).$$

(iv) The function S is convex and therefore for every convex combination of elements $x_j \in [0,1]$ we have :
$S(\sum_j \alpha_j x_j) \leq \sum_j \alpha_j \cdot S(x_j)$, where $\alpha_j \geq 0$, for all $j = 1, 2, ...,$ and $\sum_j \alpha_j = 1$.
Suppose $\alpha = \{i : m_\mu(\lambda_i) > 0\},\ \beta = \{j : m_\mu(\eta_j) > 0\},\ \Gamma = \{k : m_\mu(\gamma_k) > 0\}$ and $\delta_i = \{j : \eta_j \subseteq \lambda_i\},\ i = 1, 2, ..., n$.
Put $\alpha_j = m_\mu^+(\eta_j \mid \lambda_i),\ x_j = m_\mu^+(\gamma_k \mid \eta_j),\ i,\ k$ are fixed, and $j = 1, 2, ..., m$.
For $i \in \alpha$ we have

$$\begin{aligned}
\sum_{j\in\beta} \alpha_j &= \sum_{j\in\beta} m_\mu^+(\eta_j \mid \lambda_i) = \sum_{j=1}^{m} m_\mu^+(\eta_j \mid \lambda_i) \\
&= \sum_{j=1}^{m} \frac{m_\mu(\eta_j \cap \lambda_i)}{m_\mu(\lambda_i)} = \frac{m_\mu((\cup_{j=1}^{m} \eta_j) \cap \lambda_i)}{m_\mu(\lambda_i)} = 1.
\end{aligned}$$

Since $m_\mu(\xi \cap (\cup_{j\in\delta_i} \eta_j)) = m_\mu(\xi \cap \lambda_i)$ for every $\xi \in \mathcal{M}_\mu$, then

$$\begin{aligned}
\sum_{j\in\beta} \alpha_j\ x_j &= \sum_{j\in\beta} m_\mu^+(\eta_j \mid \lambda_i) m_\mu^+(\gamma_k \mid \eta_j) = \sum_{j\in\beta} \frac{m_\mu(\eta_j \cap \lambda_i)}{m_\mu(\lambda_i)} \cdot \frac{m_\mu(\gamma_k \cap \eta_j)}{m_\mu(\eta_j)} \\
&= \sum_{j\in\delta_i} \frac{m_\mu(\eta_j)}{m_\mu(\lambda_i)} \cdot \frac{m_\mu(\gamma_k \cap \eta_j)}{m_\mu(\eta_j)} = \frac{m_\mu(\gamma_k \cap (\cup_{j\in\delta_i} \eta_j))}{m_\mu(\lambda_i)} = \frac{m_\mu(\gamma_k \cap \lambda_i)}{m_\mu(\lambda_i)} \\
&= m_\mu^+(\gamma_k \mid \lambda_i).
\end{aligned}$$

So, we obtain

$$S(m_\mu^+(\gamma_k \mid \lambda_i)) \leq \sum_{j\in\beta} m_\mu^+(\eta_j \mid \lambda_i) S(m_\mu^+(\gamma_k \mid \eta_j)).$$

Therefore we have

$$-m_\mu(\lambda_i)S(m_\mu^+(\gamma_k \mid \lambda_i)) \geq -m_\mu(\lambda_i)\sum_{j=1}^{m} m_\mu^+(\eta_j \mid \lambda_i) S(m_\mu^+(\gamma_k \mid \eta_j))$$

for $i = 1, ..., n,\ k = 1, ..., r$. Thus

$$\begin{aligned} H(R \mid P) &= -\sum_{i=1}^{n}\sum_{k=1}^{r} m_\mu(\lambda_i)S(m_\mu^+(\gamma_k \mid \lambda_i)) \\ &\geq -\sum_{i=1}^{n}\sum_{k=1}^{r}\sum_{j=1}^{m} m_\mu(\lambda_i)m_\mu^+(\eta_j \mid \lambda_i)S(m_\mu^+(\gamma_k \mid \eta_j)) \\ &= -\sum_{j=1}^{m}\sum_{k=1}^{r} m_\mu(\eta_j)S(m_\mu^+(\gamma_k \mid \eta_j)) = H(R \mid Q). \end{aligned}$$

(v) Proposition (ii) implies that

$$H(R \vee Q \mid P) = H(Q \mid P) + H(R \mid P \vee Q).$$

Since $P \prec P \vee Q$ then (iv) implies $H(R \mid P) \geq H(R \mid P \vee Q)$. □

Theorem 3.6. *Let* $P, Q, R \in \mathbf{P}$. *Then*
(i) $H(R \mid Q) \leq H(R)$.
(ii) $H(Q \vee R) \leq H(Q) + H(R)$.
(iii) $H(R \vee Q) = H(Q) + H(R \mid Q)$.
(iv) *If* $P \prec Q$ *then* $H(P) \leq H(Q)$.

Proof. (i) Put $P = \{\mu\}$ in part (iv) of Theorem 3.5.
(ii) Let $R = \{\mu\}$. Then

$$H(P \mid R) = -\sum_{i=1}^{n} m_\mu(\mu)S(m_\mu(\lambda_i \mid \mu)) = -\sum_{i=1}^{n} S(m_\mu(\lambda_i)) = H(P).$$

Moreover if we take $P = \{\mu\}$ in part (v) of Theorem 3.5, then

$$H(Q \vee R \mid P) = H(Q \vee R),\ H(Q \mid P) = H(Q),\ H(R \mid P) = H(R).$$

(iii) Put $P = \{\mu\}$ in part (ii) of Theorem 3.5.
(iv) Put $R = \{\mu\}$ in part (iii) of Theorem 3.5. □

4. THE RELATIVE ENTROPY OF RELATIVE MEASURE PRESERVING MAPS

Let $(X, \mathcal{M}_\mu, m_\mu)$ be a $\mu-$measure space and μ be a constant observer on X, i.e. μ is a constant function.
The function $\varphi : X \longrightarrow X$ is called $\mu-$measurable if $\varphi^{-1}(\lambda) \in \mathcal{M}_\mu$ for every $\lambda \in \mathcal{M}_\mu$, where $\varphi^{-1}(\lambda)(x) = \lambda(\varphi(x)),\ \forall x \in X$. The function $\varphi : X \longrightarrow X$ is said to be a $\mu-$measure preserving map if φ is $\mu-$measurable and $m_\mu(\varphi^{-1}(\lambda)) = m_\mu(\lambda)$,

for all $\lambda \in \mathcal{M}_\mu$.

Example 4.1. *Let $X = [-1,1]$ and $\mu : X \longrightarrow [0,1]$ be defined by $\mu(x) = 7/8$ for all $x \in X$. Suppose $\lambda \subset \mu$ be defined by :*

$$\lambda(x) = \begin{cases} 7/8 & if\ x \in [-1,0) \\ 0 & if\ x \in [0,1] \end{cases}.$$

Moreover let $\mathcal{M}_\mu = \{\chi_\emptyset, \mu, \lambda, \lambda'\}$, and $m_\mu : \mathcal{M}_\mu \longrightarrow [0,1]$ be defined by : $m_\mu(\mu) = 1 = m_\mu(\lambda \cup \lambda')$, $m_\mu(\chi_\emptyset) = m_\mu(\lambda \cap \lambda') = 0$ and $m_\mu(\lambda) = m_\mu(\lambda') = 1/2$. Then $\varphi : X \longrightarrow X$ defined by $\varphi(x) = x\cos(x)$, is a $\mu-$measurable preserving transformation.

Let P be a $\mu-$measurable partition of μ. Then $\varphi^{-n}P$ is the set $\{\varphi^{-n}\lambda : \lambda \in P\}$ where $n \geq 0$.

Remark 4.2. *It is obvious that $\varphi^{-n}P$ is a $\mu-$measurable partition of μ, for each $n \geq 0$.*

Remark 4.3. *If $\varphi : X \longrightarrow X$ is a $\mu-$measure preserving transformation of $(X, \mathcal{M}_\mu, m_\mu)$ then φ^k, $k \geq 1$ is too.*
Moreover $H(\varphi^{-k}P) = H(P)$ for all $k \geq 0$.

Definition 4.4. *Let $\varphi : X \longrightarrow X$ be a $\mu-$measure preserving transformation of the relative probability space $(X, \mathcal{M}_\mu, m_\mu)$. If $P \in \mathbf{P}$ then*

$$h(\varphi, P) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \varphi^{-i}P)$$

is called the relative entropy of φ with respect to P. Moreover

$$h(\varphi) = \sup_{P\in\mathbf{P}} h(\varphi, P)$$

is called the relative entropy of φ.

Theorem 4.5. *For every $P \in \mathbf{P}$,*

$$\lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \varphi^{-i}P)$$

exists.

Proof. If $a_n = H(\bigvee_{i=0}^{n-1} \varphi^{-i}P)$ then part (ii) of Theorem 3.6, implies

$$\begin{aligned} a_{r+s} &= H(\bigvee_{i=0}^{r+s-1} \varphi^{-i}P) = H(\bigvee_{i=0}^{r-1} \varphi^{-i}P \vee \bigvee_{i=r}^{r+s-1} \varphi^{-i}P) \\ &\leq H(\bigvee_{i=0}^{r-1} \varphi^{-i}P) + H(\bigvee_{i=r}^{r+s-1} \varphi^{-i}P) = a_r + H(\varphi^{-r}(\bigvee_{i=0}^{s-1} \varphi^{-i}P)) \\ &= a_r + H(\bigvee_{i=0}^{s-1} \varphi^{-i}P) = a_r + a_s. \end{aligned}$$

So Theorem 4.9 of [7] implies that:

$$\lim_{n\to\infty}\frac{1}{n}H(\bigvee_{i=0}^{n-1}\varphi^{-i}P)$$

exists. $\square$

Theorem 4.6. *Suppose $P,Q\in\mathbf{P}$ and φ is a $\mu-$measure preserving transformation of $(X,\mathcal{M}_\mu,m_\mu)$, then*
(i) $h(\varphi,P)\le H(P)$.
(ii) $h(\varphi,P\vee Q)\le h(\varphi,P)+h(\varphi,Q)$.
(iii) If $P\prec Q$ then $h(\varphi,P)\le h(\varphi,Q)$.
(iv) $h(\varphi,P)\le h(\varphi,Q)+H(P\mid Q)$.
(v) $h(\varphi,\varphi^{-1}P)=h(\varphi,P)$.
(vi) If $k\ge 1,\ h(\varphi,P)=h(\varphi,\bigvee_{i=0}^{k}\varphi^{-i}P)$.
(vii) If φ is invertible and $k\ge 1$ then $h(\varphi,P)=h(\varphi,\bigvee_{i=-k}^{k}\varphi^{-i}P)$.

Proof. (i)

$$\begin{aligned}\frac{1}{n}H(\bigvee_{i=0}^{n-1}\varphi^{-i}P)&\le\frac{1}{n}\sum_{i=0}^{n-1}H(\varphi^{-i}P)\ by\ (ii)\ of\ Theorem\ 3.6\\&=\frac{1}{n}\sum_{i=0}^{n-1}H(P)=H(P).\end{aligned}$$

(ii)

$$\begin{aligned}H(\bigvee_{i=0}^{n-1}\varphi^{-i}(P\vee Q))&=H(\bigvee_{i=0}^{n-1}\varphi^{-i}P\vee\bigvee_{i=0}^{n-1}\varphi^{-i}Q)\\&\le H(\bigvee_{i=0}^{n-1}\varphi^{-i}P)+H(\bigvee_{i=0}^{n-1}\varphi^{-i}Q).\end{aligned}$$

(iii) If $P\prec Q$ then

$$\bigvee_{i=0}^{n-1}\varphi^{-i}P\prec\bigvee_{i=0}^{n-1}\varphi^{-i}Q,\ n\ge 1.$$

(iv)

$$\begin{aligned}H(\bigvee_{i=0}^{n-1}\varphi^{-i}P)&\le H((\bigvee_{i=0}^{n-1}\varphi^{-i}P)\vee(\bigvee_{i=0}^{n-1}\varphi^{-i}Q))\\&=H(\bigvee_{i=0}^{n-1}\varphi^{-i}Q)+H((\bigvee_{i=0}^{n-1}\varphi^{-i}P)\mid(\bigvee_{i=0}^{n-1}\varphi^{-i}Q))\end{aligned}$$

We have

$$H((\bigvee_{i=0}^{n-1}\varphi^{-i}P) \mid (\bigvee_{i=0}^{n-1}\varphi^{-i}Q)) \leq \sum_{i=0}^{n-1} H((\varphi^{-i}P) \mid (\bigvee_{j=0}^{n-1}\varphi^{-j}Q))$$
$$\leq \sum_{i=0}^{n-1} H((\varphi^{-i}P) \mid (\varphi^{-i}Q)) = nH(P \mid Q).$$

Thus

$$H(\bigvee_{i=0}^{n-1}\varphi^{-i}P) \leq H(\bigvee_{i=0}^{n-1}\varphi^{-i}Q) + nH(P \mid Q).$$

(v)

$$H(\bigvee_{i=1}^{n}\varphi^{-i}P) = H(\varphi^{-1}\bigvee_{i=0}^{n-1}\varphi^{-i}P),\ So\ h(\varphi,\varphi^{-1}P) = h(\varphi, P).$$

(vi)

$$h(\varphi, \bigvee_{i=0}^{k}\varphi^{-i}P) = \lim_{n\to\infty}\frac{1}{n}H(\bigvee_{j=0}^{n-1}\varphi^{-j}(\bigvee_{i=0}^{k}\varphi^{-i}P)) = \lim_{n\to\infty}\frac{1}{n}H(\bigvee_{i=0}^{k+n-1}\varphi^{-i}P)$$
$$= \lim_{n\to\infty}\frac{k+n}{n}\frac{1}{k+n}H(\bigvee_{i=0}^{k+n-1}\varphi^{-i}P) = h(\varphi, P).$$

(vii)

$$h(\varphi, \bigvee_{i=-k}^{k}\varphi^{-i}P) = h(\varphi, \bigvee_{i=0}^{2k}\varphi^{-i}P) = h(\varphi, P).$$

□

5. RELATIVE ENTROPY AND GENERATORS OF RELATIVE MEASURE PRESERVING MAPS

Let $(X_i, \mathcal{M}_{\mu_i}, m_{\mu_i}),\ i = 1,2$ be two relative measure spaces. Moreover let μ_i be constant observers on X_i for $i = 1,2$. The function $\psi : X_1 \longrightarrow X_2$ is called $(\mu_1,\mu_2)-$measure preserving map if:
(i) $\psi^{-1}(\eta) \in \mathcal{M}_{\mu_1}$, and,
(ii) $m_{\mu_1}(\psi^{-1}(\eta)) = m_{\mu_2}(\eta)$ for each $\eta \in \mathcal{M}_{\mu_2}$.

Example 5.1. *The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = x^3$ is a Borel-measurable with respect to usual $\sigma-$Borel Algebra of $\mathbb{R}$. Let $\mu : \mathbb{R} \longrightarrow [0,1]$ be defined by $\mu(x) = 3/4$ and Let $\gamma : \mathbb{R} \longrightarrow [0,1]$ be defined by $\gamma(x) = 1/2$. Moreover let $\mathcal{M}_\mu = \{\chi_\emptyset, \mu\}$ and $\mathcal{M}_\gamma = \{\chi_\emptyset, \gamma\}$. Then f is not $(\mu,\gamma)-$measurable, because $f^{-1}(\gamma) = \gamma$ and is not belong to $\mathcal{M}_\mu$, but if we consider $\sigma_\mu-$Algebra $\mathcal{M}_\mu = \{\chi_\emptyset, \mu, \gamma\}$ and $\mathcal{M}_\gamma = \{\chi_\emptyset, \gamma\}$, then f will be $(\mu,\gamma)-$measurable. In this case if we define $m_\mu(\chi_\emptyset) = 0,\ m_\mu(\gamma) = 1/2,\ m_\mu(\mu) = 1,\ m_\gamma(\chi_\emptyset) = 0,$ and $m_\gamma(\gamma) = 1/2$. Then f is $(\mu,\gamma)-$measure preserving map.*

Definition 5.2. *The relative dynamical system $(X_2, \mathcal{M}_{\mu_2}, m_{\mu_2}, \varphi_2)$ is a $(\mu_1, \mu_2)-$homeomorphic image of the relative dynamical system $(X_1, \mathcal{M}_{\mu_1}, m_{\mu_1}, \varphi_1)$ if there exists a $(\mu_1, \mu_2)-$measure preserving transformation $\psi : X_1 \longrightarrow X_2$ such that :*

$$\psi \circ \varphi_1 = \varphi_2 \circ \psi.$$

If ψ is invertible and $(X_1, \mathcal{M}_{\mu_1}, m_{\mu_1}, \varphi_1)$ is a $(\mu_2, \mu_1)-$ homeomorphic image of the relative dynamical system $(X_2, \mathcal{M}_{\mu_2}, m_{\mu_2}, \varphi_2)$ under the $(\mu_2, \mu_1)-$ homeomorphic transformation ψ^{-1}, we say that: $(X_i, \mathcal{M}_{\mu_i}, m_{\mu_i}, \varphi_i)$, $i = 1, 2$ are $(\mu_1, \mu_2)-$isomorphic.

The following theorem states, the relative entropy h is an $(\mu_1, \mu_2)-$isomorphism invariant.

Theorem 5.3. *If $(X_i, \mathcal{M}_{\mu_i}, m_{\mu_i}, \varphi_i)$, for $i = 1, 2$, are $(\mu_1, \mu_2)-$isomorphic with $m_{\mu_i}(\mu_i) = 1$, $i = 1, 2$ then $h(\varphi_1) = h(\varphi_2)$.*

Proof. Since $(X_2, \mathcal{M}_{\mu_2}, m_{\mu_2}, \varphi_2)$ is a $(\mu_1, \mu_2)-$ homeomorphic image of $(X_1, \mathcal{M}_{\mu_1}, m_{\mu_1}, \varphi_1)$ under an $(\mu_1, \mu_2)-$isomorphism $\psi : X_1 \longrightarrow X_2$, then $\psi \circ \varphi_1 = \varphi_2 \circ \psi$.
If Q is a μ_2-measurable partition of μ_2 where

$$\begin{aligned} h(\varphi_2, Q) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} \varphi_2^{-i} Q\Big) = \lim_{n\to\infty} \frac{1}{n} H\Big(\psi^{-1}\Big(\bigvee_{i=0}^{n-1} \varphi_2^{-i} Q\Big)\Big) \\ &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} \psi^{-1}\varphi_2^{-i} Q\Big). \end{aligned}$$

Then

$$h(\varphi_2, Q) = \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} \varphi_1^{-i}(\psi^{-1}Q)\Big) = h(\varphi_1, \psi^{-1}Q).$$

Thus

$$\sup_Q h(\varphi_2, Q) = \sup_{\psi^{-1}Q} h(\varphi_1, \psi^{-1}Q) = \sup_P h(\varphi_1, P).$$

So $h(\varphi_2) \leq h(\varphi_1)$.
Similarly we have $h(\varphi_1) \leq h(\varphi_2)$. □

In the rest of this section we also assume that μ is a constant observer on X.

Theorem 5.4. *Let φ be a $\mu-$measure preserving transformation of the relative probability space $(X, \mathcal{M}_\mu, m_\mu)$. Then*
(i) For $k > 0$, $h(\varphi^k) = kh(\varphi)$;
(ii) If φ is invertible then $h(\varphi^k) = |k|\ h(\varphi)$ for all $k \in \mathbb{Z}$.

Proof. (i) At first we show that for every $P \in \mathbf{P}$, and $k > 0$

$$h\Big(\varphi^k, \bigvee_{i=0}^{k-1} \varphi^{-i} P\Big) = kh(\varphi, P)$$

Because

$$\lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} \varphi^{-kj}\Big(\bigvee_{i=0}^{k-1} \varphi^{-i} P\Big)\Big) = \lim_{n\to\infty} \frac{k}{nk} H\Big(\bigvee_{i=0}^{nk-1} \varphi^{-i} P\Big) = kh(\varphi, P).$$

Thus

$$kh(\varphi) = k \cdot \sup_{P \in \mathbf{P}} h(\varphi, P) = \sup_{P} h(\varphi^k, \bigvee_{i=0}^{n-1} \varphi^{-i} P) \leq \sup_{Q} h(\varphi^k, Q) = h(\varphi^k).$$

Moreover $h(\varphi^k, P) \leq h(\varphi^k, \bigvee_{i=0}^{k-1} \varphi^{-i} P) = kh(\varphi, P)$ so $h(\varphi^k) \leq kh(\varphi, P)$.
(ii) We only need to show $h(\varphi^{-1}) = h(\varphi)$, and this follows from the following equalities:

$$H(\bigvee_{i=0}^{n-1} \varphi^i P) = H(\varphi^{-(n-1)} \bigvee_{i=0}^{n-1} \varphi^i P) = H(\bigvee_{j=0}^{n-1} \varphi^{-j} P).$$

□

Definition 5.5. *Let $\varphi : X \longrightarrow X$ be an invertible $\mu-$measure preserving transformation then a $\mu-$measurable partition P of μ is called a relative generator of the $(X, \mathcal{M}_\mu, m_\mu, \varphi)$ (or φ) if there exists an integer $k > 0$ such that*

$$Q \prec \bigvee_{i=0}^{k} \varphi^{-i} P$$

for every $\mu-$measurable partition Q of μ.

Theorem 5.6. *If P is a relative generator of the $\mu-$measure preserving transformation φ then $h(\varphi, Q) \leq h(\varphi, P)$, for every $\mu-$measurable partition Q of μ.*

Proof. Since P is a relative generator of φ, then for $Q \in \mathbf{P}$, there exists an integer $k > 0$ such that

$$Q \prec \bigvee_{i=0}^{k} \varphi^{-i} P.$$

Hence

$$h(\varphi, Q) \leq h(\varphi, \bigvee_{i=0}^{k} \varphi^{-i} P) = h(\varphi, P).$$

□

Theorem 5.7. *Let P be a relative generator of a $\mu-$measure preserving transformation $\varphi : X \longrightarrow X$. Then $h(\varphi) = h(\varphi, P)$.*

Proof. Theorem 5.6 implies that $h(\varphi, Q) \leq h(\varphi, P)$ for each $Q \in \mathbf{P}$.
So $\sup_{Q \in \mathbf{P}} h(\varphi, Q) = h(\varphi, P)$. Thus $h(\varphi) = h(\varphi, P)$. □

6. CONCLUSION

We show that the relative probability spaces are generalization of the soft fuzzy probability spaces. So one can use of the notion of the relative probability space to consider new concepts of the relative dynamical systems such as relative ergodicity, mixing and relative isomorphisms. Moreover deduction a method for calculating relative entropies of the $\mu-$measure preserving maps when μ is not a constant function on X, is the new subject for the future research.

References

[1] D. Markechová, The Entropy of Fuzzy Dynamical Systems and Generators, Fuzzy Sets and Systems 48 (1992) 351-363.
[2] E.P. Klement, R. Lowen and W. Schwychla, Fuzzy Probability Measure, Fuzzy Sets and Systems 5 (1981) 21-30.
[3] B. Ghazanfari, M.R. Molaei, Relative Topological Dynamics, Differential Geometry-Dynamical Systems, Vol. 8 (2006) 105-111.
[4] M.H. Anvari, M.R. Molaei, Genomic from the Viewpoint of Observational Modeling, WSEAS Transactions on Biology and Biomedicine, Issue 2, Vol. 2 (2005) 228-234.
[5] K. Piasecki, Probability of Fuzzy Events Defined As Denumerable Additivity Measure, Fuzzy Sets and Systems 17 (1985) 272-284.
[6] M.R. Molaei, Relative Semi-dynamical Systems, International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, Volume 12-2 (2004) 237-243.
[7] P. Walters, An Introduction to Ergodic Theory, Springer-Verlag, 1982.