KŌDAI MATH. SEM. REP
22 (1970), 443—468

ENTROPY AND SEMIVALUATIONS ON SEMILATTICES

By Yatsuka Nakamura

1. Introduction.

Reasoning of the popular form $-\Sigma p_i \log p_i$ of the entropy have been done by Shannon [8], Hinchin [4], Faddeev [3] and many other authors [10], [6], [5]. Their postulates of entropy would satisfy the person who is interested only in the informational feature of entropy. From the mathematical point of view, however, those are not enough to clearify how entropy is induced from the more fundamental concepts in mathematics. As one of such attempts, we shall give here a new postulates of entropy, which is, so to speak, a lattice-theoretic one.

In § 2, a simple explanation of semilattice will be given, and in § 3 a certain function on semilattice called a semivaluation will be defined, which has been used by many mathematicians on lattices. It seems that the semivaluation is, in essence, closely related to the semilattice. In the last section, a new postulates of entropy will be given: it states that a symmetric continuous function of probabilities, which gives a semivaluation on a set of all finite partitions on any probability space and independence of the semivaluation is derived from a probabilistic one, is just the entropy function.

The author wishes to express his sincere thanks to Professor H. Umegaki and Professor T. Shimogaki for many kind suggestions and encouragements in preparing this paper.

2. Semilattice.

A *semilattice* $(\Sigma, \circ)$ is a pair of a set Σ and an operation $\circ$ from $\Sigma\times\Sigma$ to Σ satisfying the following;

(S 1) $\quad x\circ x=x \quad$ for all $x\in\Sigma$,

(S 2) $\quad x\circ y=y\circ x \quad$ for all $x, y\in\Sigma$,

(S 3) $\quad x\circ(y\circ z)=(x\circ y)\circ z \quad$ for all $x, y, z\in\Sigma$.

Writing $x\leq y$ if $x\circ y=x$, we can see that $\leq$ is a partial order in Σ and the relation

$$x\circ y=\underset{(\leq)}{\text{g. l. b.}}\{x, y\}$$

Received April 9, 1970.

is satisfied. Clearly another partial order $\leq'$ can be defined dually in Σ, i.e., $x\leq' y$ if $y\circ x=y$. A semilattice considered with the partical order $\leq$ is sometimes called a *meet semilattice*, and a *join semilattice* for the order $\leq'$. If M is a partially ordered set in which g.l.b. $\{x, y\}$ exists for all $x, y\in M$, then putting $x\wedge y=$g.l.b. $\{x, y\}$ the pair $(M, \wedge)$ is a semilattice, more to speak, a meet semilattice. Dually assuming the existence of l.u.b. $\{x, y\}$ for all $x, y\in M$ and writing it as $x\vee y$, $(M, \vee)$ is a join semilattice. These facts are fully explained in [2]. A lattice is a join and meet semilattice at the same time. If $(L, \vee, \wedge)$ is a lattice with the least element O, then $(L\backslash\{O\}, \vee)$ is a join semilattice. A subset Σ' in a semilattice Σ is called a *subsemilattic* if Σ' is closed by the operation in Σ. A transformation θ from a semilattice $(\Sigma, \circ)$ into another semilattice $(\Sigma', \circ')$ is a *morphism* if

(2.1) $$\theta(x\circ y)=\theta(x)\circ'\theta(y) \qquad \text{for all} \quad x, y\in\Sigma$$

is satisfied. A morphism is called an *isomorphism* if it is bijection, an *endomorphism* if $(\Sigma, \circ)=(\Sigma', \circ')$, and an *automorphism* if it is isomorphism and endomorphism at the same time. Two semilattices are said to be *isomorphic* if there exists an isomorphism between them. If θ is a morphism between two join semilattices $(\Sigma, \vee, \leq)$ and $(\Sigma', \vee', \leq')$, then it is isotone, i.e.,

(2.2) $$x\leq y \quad \textit{implies} \quad \theta(x)\leq'\theta(y),$$

and the same fact is true in the case of meet semilattices. An isotone bijection between join (meet) semilattices with an isotone inverse is also a semilattice-isomorphism. Concerning to semilattices of finite elements, the following theorem consists.

PROPOSITION 2.1. *Let $(\Sigma, \vee)$ be a join semilattice of finite elements. L is a partially ordered set added the least element O to Σ, i.e., $L=\Sigma\cup\{O\}$ and the order $\leq$ in L is the same as in Σ for the elements in Σ and $O\leq x$ for all $x\in\Sigma$. Then $(L, \leq)$ is a lattice.*

Proof. For fixed x_0, $y_0\in L$, putting

$$B=\{x\in L;\ x\leq x_0,\ x\leq y_0\},$$

B is not empty as $O\in B$. Now we prove that there exists the greatest element in B. Putting x' as a maximal element in B and x as an arbitrary element in B, $x\vee x'=x'$ is valid from the maximality of x' and the inequality $x'\leq x\vee x'$. Hence $x\leq x'$, and so x' is the greatest element in B, which can be written $x_0\wedge y_0$, and L is actually a lattice. Q.E.D.

3. Semivaluations.

Now $(\Sigma, \circ)$ is a semilattice defined in the previous section.

DEFINITION 3.1. A real valued function $s[\cdot]$ on Σ is a semivaluation iff

$$s[x\circ y\circ z]-s[y\circ z]\leqq s[x\circ y]-s[y] \tag{3.1}$$

for every x, y and z in Σ. For a semivaluation $s[\cdot]$, we define the function $s[\cdot|\cdot]$ by

$$s[y|y']=s[y\circ y']-s[y']. \tag{3.2}$$

When the function $s[\cdot]$ is considered on a join semilattice $(\Sigma, \vee, \leq)$, the inequality (3.1) is satisfied iff

$$y\leq z \quad \text{implies} \quad s[x|z]\leqq s[x|y] \tag{3.3}$$

for every $y, z\in\Sigma$, in which case we call this a *join semivaluation.* A *meet semivaluation* is defined similarly;

$$y\leq z \quad \text{implies} \quad s[x|y]\leqq s[x|z]. \tag{3.4}$$

A semivaluation on a semilattice $(\Sigma, \circ)$ is *faithful* iff

$$s[x|y]=0 \quad \text{implies} \quad x\circ y=y. \tag{3.5}$$

LEMMA 3.1. *For a join semivaluation* $s[\cdot]$, *the followings are valid: for* $x, y, z\in\Sigma$,

(1) $\quad x\leq y$ *implies* $s[x]\leqq s[y]$,

(2) $\quad x\leq y$ *implies* $s[x|z]\leqq s[y|z]$,

and in particular if $s[\cdot]$ *is faithful, then*

(3) $\quad x\leq y$ *and* $x\neq y$ *implies* $s[x]<s[y]$.

Proof. (1): By (3.3) with $x\leq y$,

$$\begin{aligned} s[x\vee y]-s[y]&=s[x\vee y\vee y]-s[y]\leq s[x\vee y\vee x]-s[x] \\ &=s[x\vee y]-s[x] \end{aligned} \tag{3.6}$$

which implies $s[x]\leqq s[y]$. (2) is imediate from (1).

(3): If $s[x]=s[y]$, then

$$s[y|x]=s[y\vee x]-s[x]=s[y]-s[x]=0,$$

hence $y\vee x=x$ from (3.5) and $y\leq x$, which contradicts to $x\leq y$ and $x\neq y$. Q.E.D.

REMARK 1. A join semivaluation satisfying (3) of the above lemma, necessarily satisfies (3.5), and so the condition for faithfulness of join semivaluation is equivalent to the condition (3).

REMARK 2. For meet semivaluation, (1)$\sim$(3) in the lemma become dually;

(1′) $\quad x\leq y$ implies $s[x]\geq s[y]$, (2′) $$x \leq y \quad \text{implies} \quad s[x|z] \geq s[y|z],$$

if $s[\cdot]$ is faithful, then

(3′) $$x < y \quad \text{implies} \quad s[x] > s[y].$$

The more is derived for the properties of the semivaluation.

REMARK 3. Let $s[\cdot]$ be a semivaluation on a semilattice $(\Sigma, \circ)$. Then for all $x, y, z \in \Sigma$,

(1) $$s[x \circ y] = s[x] + s[y|x],$$

(2) $$s[x \circ y|z] = s[x|z] + s[y|x \circ z],$$

(3) $$s[x \circ y|z] \leqq s[x|z] + s[y|z],$$

(4) $$s[x \circ z|x] = s[z|x],$$

(5) $$s[x|z] \geqq 0.$$

Indeed, (1), (2) and (4) easily follow from the definition of semivaluation. (5) is also easely proved by Lemma 1 (3) or remark 2 (3′). (3) can be derived from (2) and the fact $s[y|x \circ z] \leqq s[y|z]$.

For a semivaluation s on a semilattice $(\Sigma, \circ)$, we introduce a new function $i[\cdot, \cdot|\cdot]$;

(3.7) $$\begin{aligned} i[x, y|z] &= s[x|z] + s[y|z] - s[x \circ y|z] \\ &= s[x|z] - s[x|z \circ y] = s[y|z] - s[y|z \circ x]. \end{aligned}$$

Now let us give a concept of independence among elements in Σ.

DEFINITION 3.2. Elements $x_1, \cdots, x_n \in \Sigma$ are said to be independent relative to an element $z \in \Sigma$ if

(3.8) $$i[x_1 \circ \cdots \circ x_k, x_{k+1}|z] = 0, \qquad k = 1, 2, \cdots, n-1$$

and denoted by

(3.9) $$x_1, \cdots, x_n \perp (z).$$

In the case of $n=2$, the condition $x_1, x_2 \perp (z)$ is same as

(3.10) $$s[x_1 \circ x_2|z] = s[x_1|z] + s[x_2|z].$$

The definition of independence is symmetric for $x_1, \cdots, x_n$. First we prove that

(3.11) $$x_1, \cdots, x_n \perp (z) \quad \text{implies} \quad x_1, \cdots, x_{k-1}, x_{k+1}, x_k, x_{k+2}, \cdots, x_n \perp (z).$$

Putting (x'_i) as the rearrangement of (x_i), then for $j < k-1$,

$$i[x_1'\circ\cdots\circ x_j', x_{j+1}'|z]=0$$

is clear, and for $j>k$,

$$x_1'\circ\cdots\circ x_j'=x_1\circ\cdots\circ x_j,\ x_{j+1}'=x_{j+1}$$

imply

$$i[x_1'\circ\cdots\circ x_j', x_{j+1}'|z]=0.$$

Moreover for $j=k-1$,

$$i[x_1\circ\cdots\circ x_{k-1}, x_{k+1}|z]\leqq i[x_1\circ\cdots\circ x_k, x_{k+1}|z]=0,$$

and for $j=k$, putting $y=x_1\circ\cdots\circ x_{k-1}$,

$$\begin{aligned} i[x_1\circ\cdots\circ x_{k-1}\circ x_{k+1}, x_k|z] &= s[y\circ x_{k+1}|z]-s[y\circ x_{k+1}|x_k\circ z] \\ &= s[y|z]+s[x_{k+1}|y\circ z]-s[y|x_k\circ z]-s[x_{k+1}|x_k\circ y\circ z] \\ &\leqq i[y, x_k|z]+s[x_{k+1}|z]-s[x_{k+1}|x_k\circ y\circ z] \\ &= i[y\circ x_k, x_{k+1}|z]=0. \end{aligned}$$

And for an arbitrary rearrangement, we can prove by repetition of the use of (3. 11). Hence we have just proved that the independence is well defined.

When the join semilattice Σ has the least element O, we call $x_1, \cdots, x_n$ being independent if $x_1, \cdots, x_n$ are independent with respect to the element O, and write

$$x_1, \cdots, x_n \perp.$$

We shall use this notation in § 6.

LEMMA 3. 2. *Let $s[\cdot]$, $s_1[\cdot]$ and $s_2[\cdot]$ be semivaluations on a semilattice Σ. Then the following facts hold:*

(1) *for an arbitrarily fixed $z\in\Sigma$, $s[\cdot|z]$ is also a semivaluation on Σ;*

(2) *for real numbers a_1, a_2 and b with $a_1\geqq 0$, $a_2\geqq 0$, $a_1s_1[\cdot]+a_2s_2[\cdot]+b$ is also a semivaluation;*

(3) *if Σ' is a subsemilattice of Σ, then the restriction of $s[\cdot]$ to Σ' is also a semivaluation on Σ'.*

The proof of this is obvious by the definition of semivaluation.

A characterization of semivaluations on lattices will be given in the following theorem:

THEOREM 3. 1. *If L is a lattice and $m[\cdot]$ is a real valued function on L, then $m[\cdot]$ is a join semivaluation iff*

i) $$m[x\wedge y]+m[x\vee y]\leqq m[x]+m[y],$$

ii) $$x\leq y \text{ implies } m[x]\leqq m[y],$$

for every x and y in L.

Proof. Let $m[\cdot]$ be a join semivaluation, then ii) is clear from lemma 1, and i) is proved by the following chain of equalities:

$$\begin{aligned} m[x]+m[y]-m[x\wedge y]-m[x\vee y] &= m[y]-m[x\wedge y]+m[x]-m[x\vee y] \\ &= m[y\vee(x\wedge y)]-m[x\wedge y]-(m[x\vee y]-m[x]) \\ &= m[y|x\wedge y]-m[y|x]\geqq 0. \end{aligned}$$

Conversely we assume that $m[\cdot]$ is a real valued function satisfying i) and ii), then for every z_1, z_2 and x in L and $z_1\leq z_2$,

$$\begin{aligned} m[x|z_1]-m[x|z_2] &= m[x\vee z_1]-m[z_1]-m[x\vee z_2]+m[z_2] \\ &= m[x\vee z_1]-m[x\vee z_2]-m[z_1]+m[z_2] \\ &= m[x\vee z_1]-m[(x\vee z_1)\vee z_2]-m[z_1]+m[z_2] \\ &\geqq m[(x\vee z_1)\wedge z_2]-m[z_2]-m[z_1]+m[z_2]\geqq 0, \end{aligned}$$

where the first inequality in the above chain is derived from i), and the second one is from ii) and by the fact

$$(x\vee z_1)\wedge z_2\geq z_1. \qquad \text{Q.E.D.}$$

REMARK 3. The dual case of the above theorem is easily obtained: $m[\cdot]$ is a meet semivaluation on L iff

(i) $\quad m[x\wedge y]+m[x\vee z]\leqq m[x]+m[y]$,

(ii) $\quad x\leq y$ implies $m[x]\geqq m[y]$.

COROLLARY 1. *If $m[\cdot]$ is a valuation on a lattice L, i.e., the conditions*

(1) $\quad m[x\wedge y]+m[x\vee y]=m[x]+m[y]$,

(2) $\quad x\leq y$ *implies* $m[x]\leqq m[y]$,

are satisfied, then $m[\cdot]$ is a join semivaluation and $-m[\cdot]$ is a meet semivaluation on L. Conversely if a function m on a lattice is a join semivaluation and $-m$ is a meet semivaluation, then m is a valuation.

Let P be a partially ordered set of finite length with O. P is called upper semimodular iff (σ) for a, b and c in P with $a\neq b$, if a and b cover c then for some $d\in P$ d covers a and b.

The following corollary tells us a relation between the semivaluation and the semimodularity.

COROLLARY 2. *A graded lattic L of finite length is upper semimodular iff its* *height function is a join semivaluation.*

The proof follows immediately from the equivalence of the condition (σ) and the inequality $h[x\wedge y]+h[x\vee y]\leqq h[x]+h[y]$ of [2].

A similar fact holds for a graded semilattice and a grade function $g[\cdot]$, i.e., g is semivaluation iff (σ) is satisfied.

The following theorem shows that a metric can be induced into a semilattice by a semivaluation, which is an extension of metric lattice.

THEOREM 3. 2. *For a semivaluation* $s[\cdot]$ *on a semilattice* Σ, *we put*

$$d(x, y)=s[x|y]+s[y|x],$$

then d *is a pseudo-metric on* Σ, *which satisfies the following relations*:

(m 1) $$d(x, y)=d(x, x\circ y)+d(x\circ y, y),$$

(m 2) $$d(x, x\circ y\circ z)=d(x, x\circ y)+d(x\circ y, x\circ y\circ z).$$

In particular, $d(\cdot, \cdot)$ *is a metric if* $s[\cdot]$ *is a faithful semivaluation.*

Proof. By remark 3. (5)

$$d(x, y)=s[x|y]+s[y|x]\geqq 0,$$

and

$$d(x, x)=2s[x|x]=0.$$

Since $s[y|x\circ z]=s[x\circ y\circ z]-s[x\circ z]\geqq 0$, the triangular inequality follows from that

$$\begin{aligned}
&d(x, y)+d(y, z)-d(x, z)\\
&=2s[x\circ y]-s[x]-s[y]+2s[y\circ z]-s[y]-s[z]-2s[x\circ z]+s[x]+s[z]\\
&=2\{s[x\circ y]-s[y]\}-2\{s[x\circ z]-s[y\circ z]\}\geqq 2\{s[x\circ y]-s[y]\}-2\{s[x\circ y\circ z]-s[y\circ z]\}\\
&=2\{s[x|y]-s[x|y\circ z]\}\geqq 0.
\end{aligned}$$

(m 1): $$\begin{aligned}d(x, x\circ y)+d(x\circ y, y)&=2s[x\circ y]-s[x]-s[x\circ y]+2s[x\circ y]-s[x\circ y]-s[y]\\
&=2s[x\circ y]-s[x]-s[y]=d(x, y),\end{aligned}$$

and (m 2) is similarly obtained. If $s[\cdot]$ is faithful, then $d(x, y)=0$ implies $s[x|y]=s[y|x]=0$ and which implies $y=x\circ y=x$. Q.E.D.

The following theorem shows that the conditions (m 1) and (m 2) in the previous theorem are sufficient for the characterization of such metric.

THEOREM 3. 3. *Let* Σ *be a join semilattice with* O *and* $d(\cdot, \cdot)$ *be a pseudo-*

metric on Σ satisfying the conditions (m 1) *and* (m 2). *Then* $s[\cdot]=d(\cdot, 0)$ *is a join semivaluation and*

$$d(x, y)=s[x|y]+s[y|x].$$

In particular, if d is metric, then $s[\cdot]$ is faithful.

Proof.

$$\begin{aligned} s[x|y]+s[y|x] &= 2s[x\circ y]-s[y]-s[x] \\ &= 2d(x\circ y, 0)-d(y, 0)-d(x, 0) \\ &= d(x\circ y, y)+d(x\circ y, x) \qquad \text{by (m 2)} \\ &= d(x, y) \qquad \text{by (m 1)}. \end{aligned}$$

The condition (3. 1) for $s[\cdot]$ is proved as follows:

$$d(x, z_1\circ z_2)-d(x, z_2)\leqq d(z_1\circ z_2, z_2) \tag{3. 12}$$

is clear from the triangular inequality, and the right hand side of (3. 12) is

$$2s[z_1\circ z_2]-s[z_2]-s[z_1\circ z_2]=s[z_1\circ z_2]-s[z_2], \tag{3. 13}$$

and the left hand of (3. 12) is

$$2s[x\circ z_1\circ z_2]-s[x]-s[z_1\circ z_2]-2s[x\circ z_2]+s[x]+s[z_2]. \tag{3. 14}$$

These (3. 12)~(3. 14) imply

$$2s[x\circ z_1\circ z_2]-s[z_1\circ z_2]-2s[x\circ z_2]+s[z_2]\leqq s[z_1\circ z_2]-s[z_2],$$

and

$$s[x\circ z_1\circ z_2]-s[z_1\circ z_2]\leqq s[x\circ z_2]-s[z_2].$$

If d is a metric, then $s[x|y]=s[x\circ y|y]=0$ implies

$$d(x\circ y, y)=s[x\circ y|y]+s[y|x\circ y]=s[x\circ y|y]=0,$$

and $x\circ y=y$. Q.E.D.

DEFINITION 3. 3. A semilattice with a (pseudeo) metric satisfying (m 1) and (m 2) is called a *(pseudo) metric semilattice.*

REMARK 4. For a join semilattice Σ with O and a semivaluation $s[\cdot]$, $s'[\cdot]=s[\cdot|O]$ is also a semivaluation and gives a same metric as one derived by s, and $s'[O]=0$. Hence we can assume $s[O]=0$ when we are mainly considering a metric on Σ. Under this assumption,

$$s[x\circ y]\leqq s[x]+s[y] \tag{3. 14}$$

is always valid for $s[z]=s[z|O]$ and by lemma 2 (3).

DEFINITION 3. 4. Let (Σ, d) and (Σ', d') be metric semilattices. We call them *isometrically isomorphic* if Σ and Σ' are semilattice-isomorphic by an isometric isomorphism ϕ, which maps from one to another.

Now let us examine some properties of metric semilattice.

LEMMA 3. 3. *Let (Σ, d) be a (pseudo) metric semilattice, and Σ' be a subsemilattice of Σ, then $(\Sigma', d|\Sigma')$ is also a (pseudo) metric semilattice.*

Proof. We can see easely that $d|\Sigma'$ (the restriction of d to Σ') satisfies (m 1) and (m 2) in Σ'.

LEMMA 3. 4. *Let (Σ, d) be a (pseudo) metric semilattice. Then the semilattice operation in Σ is uniformly continuous with the metric d.*

Proof. For any $x, y, z \in \Sigma$,

$$\Sigma' = \{x, x\circ y, z\circ x, z\circ x\circ y\}$$

is a subsemilattice of Σ. Let d' be a restriction of d to Σ. The element x is the zero element in Σ'; hence $s[\cdot]=d'(\cdot, x)$ is a semivaluation on Σ' by Th. 3.

Thus

$$\begin{aligned} d(z\circ x\circ y, z\circ x) &= d'(z\circ x\circ y, x) - d'(z\circ x, x) \\ &= s[z\circ x\circ y] - s[z\circ x] = s[x\circ y|z\circ x] \leqq s[x\circ y|x] \\ &= d'(x\circ y, x) - d'(x, x) = d(x\circ y, x). \end{aligned}$$

And similarly

$$d(z\circ y\circ x, z\circ y) \leqq d(y\circ x, y).$$

By (m 1) and (m 2),

$$\begin{aligned} d(z\circ x, z\circ y) &= d(z\circ x\circ y, z\circ x) + d(z\circ x\circ y, z\circ y) \\ &\leqq d(x\circ y, x) + d(x\circ y, y) = d(x, y). \end{aligned} \tag{3. 15}$$

Therefore

$$d(z\circ x, z\circ y) \leqq d(x, y)$$

and

$$\begin{aligned} d(z_n\circ x_n, z\circ x) &\leqq d(z_n\circ x_n, z_n\circ x) + d(z_n\circ x, z\circ x) \\ &\leqq d(x_n, x) + d(z_n, z). \end{aligned}$$

Q.E.D.

THEOREM 3. 4. *Let (Σ, d) be a metric semilattice and $(\bar{\Sigma}, \bar{d})$ be its completion with the metric d. Then $(\bar{\Sigma}, \bar{d})$ is also a metric semilattice.*

Proof. If $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in Σ, then

$$\{x_n\}=\{y_n\} \text{ in } \bar{\Sigma} \text{ iff } d(x_n, y_n)\to 0 \qquad (n\to\infty).$$

Now for $\bar{x}=\{x_n\}$ and $\bar{y}=\{y_n\}$ in $\bar{\Sigma}$, we put

$$\bar{x}\circ\bar{y}=\{x_n\circ y_n\}, \tag{3.16}$$

then the right hand side of (3.16) is a Cauchy sequence in Σ by the previous lemma. If there is another expression

$$\bar{x}=\{x'_n\} \quad \text{and} \quad \bar{y}=\{y'_n\},$$

then

$$d(x_n\circ y_n, x'_n\circ y'_n)\leqq d(x_n, x'_n)+d(y_n, y'_n)\to 0 \qquad (n\to\infty)$$

implies that $\bar{x}\circ\bar{y}$ is uniquely determined. We can see easely that the operation $\circ$ in $\bar{\Sigma}$ satisfies the conditions for semilattice. Moreover the metric $\bar{d}$ in $\bar{\Sigma}$ satisfies (m 1); in fact, for $\bar{x}=\{x_n\}$ and $\bar{y}=\{y_n\}$,

$$\begin{aligned}\bar{d}(\bar{x}, \bar{y})&=\lim_n d(x_n, y_n)\\ &\geqq \lim_n d(x_n, x_n\circ y_n)+\lim_n d(x_n\circ y_n, y_n)\\ &=\bar{d}(\bar{x}, \bar{x}\circ\bar{y})+\bar{d}(\bar{x}\circ\bar{y}, \bar{y}),\end{aligned}$$

and (m 2) is similarly proved. Q.E.D.

THEOREM 3.5. *Let (Σ, d) be a pseudo metric semilattice. Then its quotient metric space $(\mathcal{S}, D)$ is a metric semilattice.*

Proof. For $[x], [y]\in\mathcal{S}$, we define

$$[x]\circ[y]=[x\circ y], \tag{3.17}$$

then the operation is uniquely determined in $\mathcal{S}$; in fact, for $[x]=[x']$, $[y]=[y']$,

$$d(x'\circ y', x\circ y)\leqq d(x', x)+d(y', y)=0$$

by Lemma 4, hence $[x'\circ y']=[x\circ y]$.

Clearly the operation satisfies the condition for semilattices. The condition (m 1) for the metric D can be shown as following: for $[x]$, $[y]$ and $[z]$ in $\mathcal{S}$,

$$\begin{aligned}D([x], [y])&=d(x, y)=d(x, x\circ y)+d(x\circ y, y)\\ &=D([x], [x\circ y])+D([x\circ y], [y])=D([x], [x]\circ[y])+D([x]\circ[y], [y]).\end{aligned}$$

And (m 2) is similarly proved. Q.E.D.

Now we examine the relation of metrically completeness and order completeness of the metric semilattice.

LEMMA 3. 5. *Let* (Σ, d) *be a metrically complete metric join semilattice. Then* Σ *is upper conditional complete.*

Proof. Let S be a subset of Σ and x_1 be an element of which bounds S. $\mathfrak{F}$ is a class of all finite subset of S. Then, for $X \in \mathfrak{F}$,

$$\sup X \leq x_1$$

and for an arbitrarily fixed member $x_2 \in S$,

$$\sup_{X \in \mathfrak{F}} d(\sup X, x_2) \tag{3.18}$$

always exists. Let us denote the value (3. 18) by s_0.

Now it is clear that there exists a sequence $\{X_i\}_{i=1}^{\infty} \subset \mathfrak{F}$ such that

$$x_2 \in X_1, \quad X_1 \subset X_2 \subset \cdots \quad \text{and} \quad d(\sup X_n, x_2) \geqq s_0 - 2^{-n}.$$

For numbers n, m $(n > m)$,

$$\begin{aligned} d(\sup X_n, \sup X_m) &= d(\sup X_n, x_2) - d(\sup X_m, x_2) \\ &\leqq s_0 - (s_0 - 2^{-m}) = 2^{-m}, \end{aligned}$$

hence $\{\sup X_n\}$ is a Cauchy sequence in Σ. Then there exists an element $x_0 \in \Sigma$ to which $\{\sup X_n\}$ converses, and for all $x \in S$

$$\begin{aligned} d(x \vee x_0, x_0) &= \lim_n d(x \vee \sup X_n, \sup X_n) \\ &= \lim_n \{d(x \vee \sup X_n, x_2) - d(\sup X_n, x_2)\} \\ &\leqq s_0 - \lim_n d(\sup X_n, x_2) \\ &\leqq s_0 - \lim_n (s_0 - 2^{-n}) = 0, \end{aligned}$$

which shows $x \vee x_0 = x_0$ and $x \leq x_0$.

We can also show that x_0 is the least upper bound of S, because for every $y \in \Sigma$, $S \leq y$ implies $\sup X_n \leq y$ and $(\sup X_n) \vee y = y$, which shows $x_0 \vee y = y$ by the uniform continuity of the semilattice operation. Q.E.D.

4. Entropy and semivaluation.

We can find many examples of semivaluation in various fields of mathematics. As we pointed out in the previous section, a valuation on a lattice is a semivaluation; hence a measure on a Borel field is a semivaluation and a dimension of subspace in a Hilbert space is also a semivaluation. We can easily prove that a grade function on a graded semilattice satisfying the condition (σ) in the previous section is a semivaluation. An outer measure is a join semivaluation and conversely, writing an inner measure as m_*, $-m_*$ is a meet semivaluation. Let L_p^+ and L_p^- $(1 \leqq p \leqq +\infty)$ be a positive cone and a negative cone respectively of a real Banach space L_p on some measure space and $\|\cdot\|_p$ be its norm. Then the norm $\|\cdot\|_p$ is a join semivaluation on the lattice L_p^+ and a meet semivaluation on L_p^-. In the case $p<1$,

$$\|f\|_p = \left(\int f^p dm\right)^{1/p}$$

is not a norm, but we can show that $-\|\cdot\|_p$ is a meet semivaluation on a lattice of all non-negative measurable functions of which $\|\cdot\|_p$-value is finite, and a join semivaluation on the negative part.

There are many other examples of semivaluation, but the aim of this section is not to show all of them but to study the character of Shannon's entropy as a semivaluation on a lattice of all measurable finite (or countable) partitions.

Let $(X, \mathcal{X}, p)$ be a probability space and $\mathfrak{F}$ be a class of all measurable finite partitions of X. Now, for $\mathcal{A} \in \mathfrak{F}$

$$H(\mathcal{A}) = -\sum_{A \in \mathcal{A}} p(A) \log p(A) \tag{4. 1}$$

is the Shannon's entropy. $\mathfrak{F}$ is a lattice with the order of refinement $\prec$ and for $\mathcal{C}, \mathcal{D} \in \mathfrak{F}$, $\mathcal{C} \vee \mathcal{D}$ is represented as

$$\mathcal{C} \vee \mathcal{D} = \{C \cap D;\ C \in \mathcal{C},\ D \in \mathcal{D}\},$$

then as well known in a measure theoretic information theory (cf. [1]), for $\mathcal{A}, \mathcal{B}, \mathcal{B}' \in \mathfrak{F}$

$$\mathcal{B}' \prec \mathcal{B} \text{ implies } H(\mathcal{A} \vee \mathcal{B}) - H(\mathcal{B}) \leqq H(\mathcal{A} \vee \mathcal{B}') - H(\mathcal{B}'). \tag{4. 2}$$

The condition (4. 2) is just the condition (3. 1) of join semivaluation. Denoting $H(\mathcal{E} | \mathcal{F}) = H(\mathcal{E} \vee \mathcal{F}) - H(\mathcal{F})$, (4. 2) is written

$$\mathcal{B}' \prec \mathcal{B} \text{ implies } H(\mathcal{A} | \mathcal{B}) \leqq H(\mathcal{A} | \mathcal{B}'). \tag{4. 3}$$

The well-known equalities and inequalities for the entropy are derived from the fact that the entropy is a semivaluation; for example, by lemma 3. 1 and lemma 3. 2,

(i) $\mathcal{A} \prec \mathcal{B}$ implies $H(\mathcal{A}) \leqq H(\mathcal{B})$

and $H(\mathcal{A} \mid \mathcal{C}) \leqq H(\mathcal{B} \mid \mathcal{C}) \quad (\mathcal{C} \in \mathfrak{F})$,

(ii) $$H(\mathcal{A} \vee \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B} | \mathcal{A}),$$

$$H(\mathcal{A} \vee \mathcal{B} | \mathcal{C}) = H(\mathcal{A} | \mathcal{C}) + H(\mathcal{B} | \mathcal{C} \vee \mathcal{A}),$$

(iii) $$H(\mathcal{A} \vee \mathcal{B} | \mathcal{C}) \leqq H(\mathcal{A} | \mathcal{C}) + H(\mathcal{B} | \mathcal{C})$$

and

$$H(\mathcal{A} \vee \mathcal{B}) \leqq H(\mathcal{A}) + H(\mathcal{B})$$

because $H(O)=0$, where O is the trivial partition of X, i.e., $O=\{X\}$.

The partitions $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n \in \mathfrak{F}$ are mutually independent in the sence of § 3, i.e., $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n \perp$ if and only if they are mutually probabilistically independent. An i-function defined by (3. 7) corresponds to the mutual information

$$I(\mathcal{A}, \mathcal{B}) = \sum_A \sum_B p(A \cap B) \log \frac{p(A \cap B)}{p(A)p(B)}$$

$$= H(\mathcal{A}) + H(\mathcal{B}) - H(\mathcal{A} \vee \mathcal{B}).$$

A metric induced by the semivaluation of entropy is one which has been called the entropy metric;

$$\rho(\mathcal{A}, \mathcal{B}) = H(\mathcal{A} | \mathcal{B}) + H(\mathcal{B} | \mathcal{A}). \tag{4. 4}$$

The entropy is not faithful as the smivaluation on $\mathfrak{F}$, but of course the entropy is faithful on the quotient metric space $\mathfrak{F}/\sim$ in the sense of Th. 3. 5. And $\mathfrak{F}/\sim$ is just the classification of $\mathfrak{F}$ identifying the probalilistically same partitions. Generally $\mathfrak{F}$ (or $\mathfrak{F}/\sim$) is not necessarily complete with respect to the entropy metric, and we can complete them as in Th. 3. 4. Let us show in the following theorem that, after the completion, the generalized semivaluation can be represented by the von Neumann's entropy, i.e., for a countable (or finite) partition $\mathcal{A} = \{A_1, A_2, \cdots\}$

$$\overline{H}(\mathcal{A}) = -\sum_{i=1}^{\infty} p(A_i) \log p(A_i). \tag{4. 5}$$

Let $\mathfrak{Z}$ be a set of all finite or countable partitions with finite von Neumann's entropy. Then $\mathfrak{F} \subset \mathfrak{Z}$, and the restriction of $\overline{H}$ on $\mathfrak{F}$ is the Shannon's entropy H. We can see easely that $\mathfrak{Z}$ is a lattice and $\overline{H}$ is also a join semivaluation on $\mathfrak{Z}$. We write

$$\overline{H}(\mathcal{E} | \mathcal{F}) = \overline{H}(\mathcal{E} \vee \mathcal{F}) - \overline{H}(\mathcal{F}).$$

Write

$$\bar{\rho}(\mathcal{A}, \mathcal{B}) = \overline{H}(\mathcal{A} | \mathcal{B}) + \overline{H}(\mathcal{B} | \mathcal{A}), \qquad \mathcal{A}, \mathcal{B} \in \mathfrak{Z}. \tag{4. 6}$$

Then as the case ρ for $\mathfrak{F}$, the function $\bar{\rho}$ defines a metric over the family $\mathfrak{Z}$, and it is an extension of ρ over $\mathfrak{F}$ onto $\mathfrak{Z}$. Now we can prove the following:

THEOREM 4. 1.[1)] *$\mathfrak{Z}$ is a complete metric semilattice with the metric $\bar{\rho}$ and it is the completion of $\mathfrak{F}$ with metric ρ, that is, $\mathfrak{Z}$ is the smallest complete metric semilattice containing $\mathfrak{F}$.*

---

1) The result of similar type was previously proved by Rokhlin [7] in a case of abstract Lebesgue space.

Before proving this, let us see the following proposition, which is known in more general type.

PROPOSITION 4.1. *Let $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2, \cdots$ be partitions in $\mathfrak{Z}$ satisfying $\mathcal{B}_1 \succ \mathcal{B}_2 \succ \cdots$ and $\mathcal{B} = \bigcap_{n=1}^{\infty} \mathcal{B}_n$. Then*

$$\text{(i)} \qquad \lim_n \bar{H}(\mathcal{C} \mid \mathcal{B}_n) = \bar{H}(\mathcal{C} \mid \mathcal{B}) \qquad (\mathcal{C} \in \mathfrak{Z}),$$

$$\text{(ii)} \qquad \lim_n \bar{H}(\mathcal{B}_n) = \bar{H}(\mathcal{B}),$$

And if $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2, \cdots$ are $\mathcal{B}_1 \prec \mathcal{B}_2 \prec \cdots$ and $\mathcal{B} = \bigvee_{n=1}^{\infty} \mathcal{B}_n$, then

$$\text{(iii)} \qquad \lim_n \bar{H}(\mathcal{C} \mid \mathcal{B}_n) = \bar{H}(\mathcal{C} \mid \mathcal{B}) \qquad (\mathcal{C} \in \mathfrak{Z}),$$

$$\text{(iv)} \qquad \lim_n \bar{H}(\mathcal{B}_n) = \bar{H}(\mathcal{B}).$$

Proof. (i): Put $f(t) = -t \log t$ $(t>0)$, $=0$ $(t=0)$. Combining both the martingale and Lebesgue's convergence theorems, we have

$$\lim_n \int f(p(C \mid \mathcal{B}_n))dp = \int f(p(C \mid \mathcal{B}))dp, \qquad C \in \mathcal{C}. \tag{4.7}$$

And by the monoton convergence theorem,

$$\int \sum_{C \in \mathcal{C}} f(p(C \mid \mathcal{B}_n))dp = \sum_{C \in \mathcal{C}} \int f(p(C \mid \mathcal{B}_n))dp. \tag{4.8}$$

Moreover, by the Jensen's inequality,

$$\int f(p(C \mid \mathcal{B}_1))dp \leqq \int f(p(C \mid \mathcal{B}_2))dp \leqq \cdots \leqq \int f(p(C \mid \mathcal{B}))dp. \tag{4.9}$$

Hence by (4.8)

$$\begin{aligned} \lim_n \bar{H}(\mathcal{C} \mid \mathcal{B}_n) &= \lim_n \sum_{C \in \mathcal{C}} \int f(p(C \mid \mathcal{B}_n))dp \\ &= \sum_{C \in \mathcal{C}} \lim_n \int f(p(C \mid \mathcal{B}_n))dp \\ &= \sum_{C \in \mathcal{C}} \int f(p(C \mid \mathcal{B}))dp = \bar{H}(\mathcal{C} \mid \mathcal{B}). \end{aligned}$$

(ii): By the result of (i),

$$\lim_n \bar{H}(\mathcal{B}_1 \mid \mathcal{B}_n) = \bar{H}(\mathcal{B}_1 \mid \mathcal{B}),$$

which implies

$$\bar{H}(\mathcal{B}_n)=\bar{H}(\mathcal{B}_1)-\bar{H}(\mathcal{B}_1|\mathcal{B}_n)\to\bar{H}(\mathcal{B}_1)-\bar{H}(\mathcal{B}_1|\mathcal{B})=\bar{H}(\mathcal{B}).$$

(iii) and (iv) follow quite similarly to the proofs of (i) and (ii) respectively.

Proof of the theorem. Let $\{\mathcal{A}_n\}\subset\mathfrak{Z}$ be a Cauchy sequence with respect to the the induces metric $\bar{\rho}$ in $\mathfrak{Z}$. Then we can assume that

$$\bar{\rho}(\mathcal{A}_n, \mathcal{A}_{n+1})<1/2^n \qquad (n=1, 2, \cdots),$$

and let us find out the limit of $\mathcal{A}_n$ in $\mathfrak{Z}$. Putting

$$\mathcal{B}_{n,r}=\mathcal{A}_n\vee\cdots\vee\mathcal{A}_{n+r},$$

we see

$$\mathcal{B}_{n,1}\prec\mathcal{B}_{n,2}\prec\cdots \qquad (n=1, 2, \cdots),$$

and

$$\begin{aligned}\bar{H}(\mathcal{B}_{n,r+1})-\bar{H}(\mathcal{B}_{n,r})&=\bar{H}(\mathcal{A}_{n+r+1}|\mathcal{A}_n\vee\cdots\vee\mathcal{A}_{n+r})\\ &\leqq\bar{H}(\mathcal{A}_{n+r+1}|\mathcal{A}_{n+r})\leqq\bar{\rho}(\mathcal{A}_{n+r}, \mathcal{A}_{n+r+1})<1/2^{n+r}.\end{aligned}$$

Hence

$$\lim_r \bar{H}(\mathcal{B}_{n,r})<\infty$$

and $\mathcal{C}_n=\vee_{r=1}^{\infty}\mathcal{B}_{n,r}$ $(n=1, 2, \cdots)$ is a decreasing sequence in $\mathfrak{Z}$.[2] Writing $\mathcal{C}=\wedge_n\mathcal{C}_n$, and using the proposition, we see

$$\lim_n \bar{H}(\mathcal{C}_n)=\bar{H}(\mathcal{C}).$$

Then

$$\begin{aligned}\bar{\rho}(\mathcal{C}_n, \mathcal{C})&=\bar{H}(\mathcal{C}|\mathcal{C}_n)+\bar{H}(\mathcal{C}_n|\mathcal{C})=\bar{H}(\mathcal{C}_n|\mathcal{C})\\ &=\bar{H}(\mathcal{C}_n)-\bar{H}(\mathcal{C})\to 0 \qquad (\text{as } n\to\infty),\end{aligned}$$

and

$$\lim_r \bar{\rho}(\mathcal{B}_{n,r}, \mathcal{C}_n)=\lim_r \bar{H}(\mathcal{C}_n|\mathcal{B}_{n,r})=0$$

implies

$$\begin{aligned}\bar{\rho}(\mathcal{A}_n, \mathcal{C}_n)=\bar{\rho}\left(\mathcal{A}_n, \bigvee_r \mathcal{B}_{n,r}\right)&\leqq\sum_{r=1}^{\infty}\bar{\rho}(\mathcal{B}_{n,r+1}, \mathcal{B}_{n,r})\\ &\leqq 1/2_n\to 0 \qquad (\text{as } n\to\infty).\end{aligned}$$

Therefore

$$\bar{\rho}(\mathcal{A}_n, \mathcal{C})\leqq\bar{\rho}(\mathcal{A}_n, \mathcal{C}_n)+\bar{\rho}(\mathcal{C}_n, \mathcal{C})\to 0 \qquad (\text{as } n\to\infty),$$

2) For every finite subfield $\mathcal{B}\in\mathfrak{F}$ of $\mathcal{C}_n$, and every $\varepsilon>0$, there exists a number r_0 and a finite subfield $\mathcal{B}'$ of $\mathcal{B}_{n,r_0}$ such that $|\mathcal{B}, \mathcal{B}'|<\varepsilon$. Therefore $\bar{H}(\mathcal{B})\leqq\lim_r \bar{H}(\mathcal{B}_{n,r})<\infty$ for every $\mathcal{B}\subset\mathcal{C}_n$, which implies that $\mathcal{C}_n$ is essentially a countable partition.

which shows that C is the limit element of $\mathcal{A}_n$ in $\mathfrak{Z}$.

$\mathfrak{Z}$ is completion of $\mathfrak{F}$. In fact, for an arbitrary countable partition $\mathcal{A}=\{A_1, A_2, \cdots\}\in\mathfrak{Z}$ the sequence

$$\mathcal{A}_n=\left\{A_1, A_2, \cdots, A_{n-1}, \bigcup_{j\geqq n} A_j\right\}\in\mathfrak{F} \qquad (n=1, 2, \cdots)$$

makes a Cauchy sequence converging to $\mathcal{A}$ in $\mathfrak{Z}$. Q.E.D.

5. Abstract dynamical system.

Let (Σ, s, τ) be a triple of a semilattice Σ, a semivaluation $s=s[\cdot]$ on Σ and an endomorphism τ on Σ with which s is invariant, i.e., $s[x]=s[\tau x]$ for all $x\in\Sigma$.

We will call the triple (Σ, s, τ) as an *abstract dynamical system* (*AD-system*).

DEFINITION 5.1. Two AD-systems (Σ, s, τ) and $(\Sigma', s' \tau')$ are said to be *isomorphic* if there exists a semilattice isomorphism from Σ to Σ' such that

(a) $$s[x]=s'[\theta x], \qquad x\in\Sigma,$$

(b) $$\theta(\tau x)=\tau'(\theta x), \qquad x\in\Sigma,$$

are satisfied. And they are *weakly isomorphic* if there exists a morphism θ_1 from Σ to Σ' and another morphism θ_2 from Σ' to Σ such that

(a′) $$s[x]=s'[\theta_1 x], \qquad x\in\Sigma,$$

$$s'[y]=s[\theta_2 y], \qquad y\in\Sigma',$$

(b′) $$\theta_1(\tau x)=\tau'(\theta_1 x), \qquad x\in\Sigma,$$

$$\theta_2(\tau' y)=\tau(\theta_2 y), \qquad y\in\Sigma'.$$

Clearly the isomorphic implies the weakly isomorphic. The above definition is a generalization of the concept of the isomorphism and Sinai's weakly isomorphism of the usual (probabilistic) dynamical system.

DEFINITION 5.2. A *mean transfer value* $t(\tau, x; s)$ *of* x *with* τ is the following limit:

(5.1) $$t(\tau, x; s)=\lim_n \frac{1}{n} s[x\circ\tau x\circ\cdots\circ\tau^{n-1}x],$$

where the limit always exists as

$$\begin{aligned} d_n &= s[x\circ\tau x\circ\cdots\circ\tau^n x]-s[x\circ\tau x\circ\cdots\circ\tau^{n-1}x] \\ &= s[x\circ\tau x\circ\cdots\circ\tau^n x]-s[\tau x\circ\cdots\circ\tau^n x] \\ &= s[x|\tau x\circ\cdots\circ\tau^n x] \end{aligned}$$

is monoton decreasing (as n increase), then d_n converges to some limit d and putting $d_0=s[x]$,

$$t(\tau, x; s)=\lim_n \frac{1}{n}\sum_{k=0}^{n-1} d_k=d,$$

which can be rewritten

$$t(\tau, x; s)=\lim_n s[x|\tau x\circ\cdots\circ\tau^n x]. \tag{5. 2}$$

The property of the mean transfer value is studied in the following lemma.

LEMMA 5. 1. *Let (Σ, s, τ) be an AD-system and $t(\tau, x; s)$ be the mean transfer value of x with τ, then*

(i) $\quad x\circ y=y$ *implies* $t(\tau, x; s)\leqq t(\tau, y; s)$,

(ii) *if τ is an automorphism, then*

$$t(\tau, x; s)=t(\tau^{-1}, x; s),$$

(iii) $$t(\tau, x; s)\leqq t(\tau, y; s)+s[x|y],$$

(iv) $$|t(\tau, x; s)-t(\tau, y; s)|\leqq d(x, y),$$

where the metric of the right hand side is the one induced by the semivaluation s.

Proof. (i): Clear by the isotonicity of s.
(ii): (5. 1) and

$$s[x\circ\cdots\circ\tau^{n-1}x]=s[\tau^{-n+1}(x\circ\cdots\circ\tau^{n-1}x)]=s[x\circ\cdots\circ\tau^{-n+1}x]$$

show the result.
(iii): We can see by the following chain of formulae.

$$\frac{1}{n}s[x\circ\cdots\circ\tau^{n-1}x]\leqq\frac{1}{n}s[y\circ\cdots\circ\tau^{n-1}y\circ x\circ\cdots\circ\tau^{n-1}x]$$

$$=\frac{1}{n}s[y\circ\cdots\circ\tau^{n-1}y]+\frac{1}{n}s[x\circ\cdots\circ\tau^{n-1}x|y\circ\cdots\circ\tau^{n-1}y]$$

$$\leqq\frac{1}{n}s[y\circ\cdots\circ\tau^{n-1}y]+\frac{1}{n}\sum_{i=0}^{n-1}s[\tau^i x|y\circ\cdots\circ\tau^{n-1}y]$$

$$\leqq\frac{1}{n}s[y\circ\cdots\circ\tau^{n-1}y]+\frac{1}{n}\sum_{i=0}^{n-1}s[\tau^i x|\tau^i y]$$

$$=\frac{1}{n}s[y\circ\cdots\circ\tau^{n-1}y]+s[x|y],$$

where the last equality follows from

$$S[\tau^i x|\tau^i y]=s[x|y] \qquad (i=0, \cdots, n-1).$$

(iv): By (iii),

$$t(\tau, x; s)-t(\tau, y; s)\leqq s[x|y]\leqq d(x, y),$$

and

$$t(\tau, y; s)-t(\tau, x; s)\leqq d(x, y),$$

therefore

$$|t(\tau, x; s)-t(\tau, y; s)|\leqq d(x, y). \qquad \text{Q.E.D.}$$

DEFINITION 5. 3. A *transfer value* $t(\tau; s)$ *of* τ is defined by

$$t(\tau; s)=\sup_{x\in\Sigma} t(\tau, x; s), \tag{5. 3}$$

which may take the value $+\infty$.

Then we can see that the transfer value is an invariant of isomorphism between the abstract dynamical systems.

THEOREM 5. 1. *If two AD-systems* (Σ, s, τ) *and* (Σ', s', τ') *are weakly isomorphic, then*

$$t(\tau; s)=t(\tau', s'). \tag{5. 4}$$

Proof. Let us denote the morphisms as

$$\theta_1\colon \Sigma\to\Sigma',$$

$$\theta_2\colon \Sigma'\to\Sigma,$$

which make the systems weakly isomorphic. Then for an arbitrary $x\in\Sigma$,

$$\begin{aligned} t(\tau'; s')&\geqq t(\tau', \theta_1 x; s') \\ &=\lim_n \frac{1}{n} s'[\theta_1 x\circ\tau'\theta_1 x\circ\cdots\circ\tau'^{n-1}\theta_1 x] \\ &=\lim_n \frac{1}{n} s'[\theta_1 x\circ\theta_1\tau x\circ\cdots\circ\tau'^{n-2}\theta_1\tau x] \\ &=\lim_n \frac{1}{n} s'[\theta_1(x\circ\tau x\circ\cdots\circ\tau^{n-1}x)] \\ &=\lim_n \frac{1}{n} s[x\circ\tau x\circ\cdots\circ\tau^{n-1}x] \\ &=t(\tau, x; s), \end{aligned}$$

which shows that

$$t(\tau'; s') \geqq t(\tau; s),$$

and the converse inequality is symmetrically derived. Q.E.D.

COROLLARY. *If there exists only a morphism θ_1 from Σ to Σ' satistying the equalities for θ_1 in* (a') *and* (b'), *then*

$$t(\tau'; s') \geqq t(\tau; s).$$

The following lemma gives a method to calculate the transfer value for some examples of AD-systems.

LEMMA 5. 2. *If there exists a sequence $\{x_n\}$ in Σ, such that*

$$\lim_n s[x|x_n] = 0 \qquad \textit{for all} \quad x \in \Sigma,$$

then

$$t(\tau; s) = \varliminf_n t(\tau, x_n; s).$$

Proof. By Lemma 1 (iii),

$$t(\tau, x; s) \leqq t(\tau, x_n; s) + s[x|x_n].$$

Hence

$$t(\tau, x; s) \leqq \varliminf_n t(\tau, x_n; s),$$

and as $x \in \Sigma$ is arbitrary

$$t(\tau; s) \leqq \varliminf_n t(\tau, x_n; s)$$

is valid. And the converse inequality is clear. Q.E.D.

LEMMA 5. 3. *Let (Σ, s, τ) be an AD-system, where τ is an automorphism on Σ, then,*

$$t(\tau^n; s) = nt(\tau; s), \qquad n = 1, 2, \cdots. \tag{5.5}$$

Proof. For every integer $n \geqq 1$ and $x \in \Sigma$,

$$\begin{aligned} t(\tau^n, x; s) &= \lim_k \frac{1}{k} s[x \circ \tau^n x \circ \cdots \circ \tau^{n(k-1)} x] \\ &\leqq \lim_k n \cdot \frac{1}{nk} s[x \circ \tau x \circ \cdots \circ \tau^{nk-1} x] \\ &= nt(\tau, x; s), \end{aligned}$$

hence

$$t(\tau^n; s) \leqq nt(\tau; s).$$

To prove the converse inequality, let us write

$$y = x \circ \tau x \circ \cdots \circ \tau^{n-1} x,$$

then

$$t(\tau^n; s) \geqq t(\tau^n, y; s) = \lim_k \frac{1}{k} s[y \circ \tau^n y \circ \cdots \circ \tau^{n(k-1)} y]$$

$$= \lim_k \frac{1}{k} s[x \circ \tau x \circ \cdots \circ \tau^{nk-1} x] = nt(\tau, x; s),$$

which shows

$$t(\tau^n; s) \geqq nt(\tau; s).$$

For a negative integer n, (5.5) is also true by lemma 1 (ii). Q.E.D.

As in § 4, let $(X, \mathfrak{X}, p)$ be a probability measure space and $\mathfrak{F}$ be a class of all measurable finite partitions of X. Now T is a measure preserving transformation on X, then T^{-1} can be seen as a semilattice morphism from $\mathfrak{F}$ to $\mathfrak{F}$ itself, which preserves the Shannon's entropy H invariant.

The mean transfer value $t(T^{-1}, \mathcal{A}; H)$ of $\mathcal{A} \in \mathfrak{F}$ with the entropy H in this case is just the entropy of the measure preserving transformation T relative to a finite partition $\mathcal{A}$, and which is denoted by $h(T, \mathcal{A})$. And the transfer value $t(T^{-1}; H)$ of T^{-1} is the (Kolmogorov's) entropy $h(T)$ of T.

Other examples of the mean transfer value and the transfer value can be given in many directions. Let $(\Omega, \mathcal{S}, m)$ be an infinite measure space and S be a measure preserving transformation on Ω. Then for a finite measurable subset $A \in \mathcal{S}$ $(m(A) < \infty)$, the mean transfer value is written

$$t(s^{-1}, A; m) = \lim_n \frac{1}{n} m(A \cup \cdots \cup s^{-n+1} A),$$

and the transfer value $t(s^{-1}; m)$ is also defined. For an isomorphism problem of such infinite dynamical systems, we can use the above transfer value as an invariant for the isomorphism.

9. A characterization of entropy.

Let Δ_n be a fundamental simplex in n-dimensional Euclidean space R^n, i.e.,

$$\Delta_n = \left\{ (p_1, \cdots, p_n) \in R^n;\ p_i \geqq 0\ (i = 1, 2, \cdots, n),\ \sum_{i=1}^n p_i = 1 \right\}.$$

Now, let us consider a class of functions $\{f_n\}_{n=1}^{\infty}$ such that each f_n is defined and real valued on Δ_n and satisfies:

I. f_n is continuous,

II. if π is any parmutation of numbers $(1, 2, \cdots, n)$, then

$$f_n(p_1, \cdots, p_n) = f_n(p_{\pi(1)}, \cdots, p_{\pi(n)}),$$

III. $$f_{n+1}(p_1, \cdots, p_n, 0) = f_n(p_1, \cdots, p_n).$$

For an arbitrary probability space $(X, \mathcal{X}, p)$ and a class $\mathfrak{F}$ of measurable finite partitions of X, we can define a function F on $\mathfrak{F}$ using the above functions f_n as

$$F(\mathcal{A}) = f_n(p(A_1), \cdots, p(A_n)),$$

where $\mathcal{A} = \{A_1, \cdots, A_n\} \in \mathfrak{F}$. Then F is well defined on $\mathfrak{F}$ by the conditions I$\sim$III.

Further let us consider the following two conditions concerning the function F:

IV. F is a join semivaluation on $\mathfrak{F}$, i.e.,

$$F(\mathcal{A} \vee \mathcal{B}) + F(\mathcal{A} \wedge \mathcal{B}) \leqq F(\mathcal{A}) + F(\mathcal{B}),$$

and

$$\mathcal{C} \prec \mathcal{D} \Rightarrow F(\mathcal{C}) \leqq F(\mathcal{D}),$$

V. if $\mathcal{A} \in \mathfrak{F}$ and $\mathcal{B} \in \mathfrak{F}$ are probabilistically independent, then they are also independent with the semivaluation F, i.e.,

$$F(\mathcal{A} \vee \mathcal{B}) = F(\mathcal{A}) + F(\mathcal{B}).$$

We can easily show an example of a class of functions $\{f_n\}$ satisfying I$\sim$V. In fact the entropy function

$$f_n(p_1, \cdots, p_n) = -c \sum_{i=1}^{n} p_i \log p_i \qquad (c\text{: constant}) \tag{6.1}$$

satisfies the conditions. Do there exist other such functions? The aim of this section is to show that such function is only of the entropy type. Therefore we can employ the conditions I$\sim$V for the axioms of the entropy.

THEOREM 6.1. *A class of functions $\{f_n\}$ satisfying* I$\sim$V *is written by the form* (6.1).

Before proving the theorem let us see some lemmas and propositions. Take a probability measure space $(X, \mathcal{X}, p)$, and F is the function on $\mathfrak{F}$ defined in the above paragraph, i.e., defined by any class of functions $\{f_n\}$ satisfying I$\sim$V.

LEMMA 6.1. *If the partitions $\mathcal{B}_n \in \mathfrak{F}$ $(n=1, 2, \cdots)$ satisfy $\mathcal{B}_n \prec \mathcal{B}_{n+1}$ $(n=1, 2, \cdots)$ and $\vee_{n=1}^{\infty} \mathcal{B}_n = \mathcal{X}$, then for any $\mathcal{A} \in \mathfrak{F}$,*

$$\lim_n F(\mathcal{A} | \mathcal{B}_n) = 0.$$

Proof. Writing $\mathcal{A}=\{A_1, \cdots, A_m\}$, for any number $k>0$ there exists a number $n_0=n_0(k)$ and

$$p(A_i \triangle B_i)<\frac{1}{2mk} \qquad (i=1, \cdots, m)$$

for some $\mathcal{B}_{n_0}$-measurable sets B_i $(i=1, \cdots, m)$. As $\{B_1, \cdots, B_m\}$ is not a partition in general, we put

$$\bar{B}_i=B_i \setminus \bigcup_{j<i} B_j \qquad (i=1, \cdots, m),$$

$$\bar{B}_{m+1}=X \setminus \bigcup_i B_i,$$

then clearly $\mathcal{B}^{(k)}=\{\bar{B}_1, \cdots, \bar{B}_{m+1}\}$ is a measurable finite partition. And from

$$B_i \cap B_j \subset (A_i \cap A_j) \cup (A_i \triangle B_i) \cup (A_j \triangle B_j),$$

we can see that

$$p(A_i \triangle \bar{B}_i) \leqq p(A_i \triangle B_i)+p(B_i \triangle \bar{B}_i)$$

$$\leqq \frac{1}{2mk}+\sum_{j<i} p(B_i \cap B_j)<\frac{1}{k} \qquad (i=1, \cdots, m).$$

Hence

$$|p(A_i)-p(\bar{B}_i)| \leqq p(A_i \triangle \bar{B}_i)<\frac{1}{k} \qquad (i=1, \cdots, m). \tag{6. 2}$$

And

$$p(\bar{B}_{m+1})=p\Big(X \setminus \bigcup_i B_i\Big) \leqq p\Big(\bigcup_i A_i \triangle B_i\Big)<\frac{1}{k},$$

which and (6. 2) imply

$$|F(\mathcal{B}^{(k)})-F(\mathcal{A})| \rightarrow 0 \qquad \text{as} \quad k \rightarrow \infty. \tag{6. 3}$$

On the other hand

$$|p(A_i)-p(A_i \cap \bar{B}_i)| \leqq p(A_i \triangle \bar{B}_i)<\frac{1}{k},$$

$$p(A_i \cap \bar{B}_j)<p(A_j \triangle \bar{B}_j)<\frac{1}{k} \qquad (i \neq j)$$

and

$$p(A_i \cap \bar{B}_{m+1}) \leqq p(\bar{B}_{m+1})<\frac{1}{k}$$

imply

$$
\begin{aligned}
F(\mathcal{A}\vee\mathcal{B}^{(k)}) &= f_{m(m+1)}(\{p(A_i\cap\bar{B}_j);\ i=1,\cdots,m,\ j=1,\cdots,m+1\}) \\
&\underset{(k\to\infty)}{\longrightarrow} f_{m(m+1)}(\{p(A_i)\delta_{ij};\ i=1,\cdots,m,\ j=1,\cdots,m+1)\} \\
&= f_m(p(A_1),\cdots,p(A_m)) = F(\mathcal{A}).
\end{aligned}
\tag{6. 4}
$$

Therefore, for any number $l>n_0(k)$

$$
\begin{aligned}
F(\mathcal{A}\,|\,\mathcal{B}_l) &\leqq F(\mathcal{A}\,|\,\mathcal{B}^{(k)}) = F(\mathcal{A}\vee B^{(k)}) - F(\mathcal{B}^{(k)}) \\
&= \{F(\mathcal{A}\vee\mathcal{B}^{(k)}) - F(\mathcal{A})\} + \{F(\mathcal{A}) - F(\mathcal{B}^{(k)})\}.
\end{aligned}
\tag{6. 5}
$$

For every $\varepsilon>0$ we can choose k so large that the last terms of (6. 5) are smaller than ε by (6. 3) and (6. 4), i.e., for any $\varepsilon>0$ there exists a number n_0 and

$$F(\mathcal{A}\,|\,\mathcal{B}_l)<\varepsilon \qquad (l>n_0)$$

is satisfied. Q.E.D.

Now, let us consider an invertible measure preserving transformation T on $(X,\mathcal{X},p)$, then we can construct the mean transfer value $t(T^{-1},\mathcal{A};F)$ of $\mathcal{A}\in\mathfrak{F}$ with T^{-1} and the transfer value $t(T^{-1};F)$ of T^{-1}. The following proposition was proved by Kolmogorov and Sinai in the case of the entropy function.

PROPOSITION 6. 1. *If T is invertible and $\vee_{k=-\infty}^{\infty}T^k\mathcal{A}=\mathcal{X}$, then*

$$t(T^{-1};F)=t(T^{-1},\mathcal{A};F).$$

Proof. Let us denote $\mathcal{B}_n=\vee_{k=-n}^{n}T^k\mathcal{A}$ $(n=1,2,\cdots)$, then the sequence $\{\mathcal{B}_n\}$ satisfies the assumptions of Lemma 6. 1. Hence

$$\lim_n F(C\,|\,\mathcal{B}_n)=0 \qquad \text{for all} \quad C\in\mathfrak{F},$$

and so

$$t(T^{-1};F)=\varliminf_n t(T^{-1},\mathcal{B}_n;F)$$

by Lemma 5. 2. Moreover

$$
\begin{aligned}
t(T^{-1},\mathcal{B}_n;F) &= \lim_k \frac{1}{k}F(\mathcal{B}_n\vee T^{-1}\mathcal{B}_n\vee\cdots\vee T^{-k+1}\mathcal{B}_n) \\
&= \lim_k \frac{1}{k}F\left(\bigvee_{l=-n-k+1}^{n} T^l\mathcal{A}\right) = \lim_k \frac{1}{k}F\left(\bigvee_{l=-k+1}^{2n} T^l\mathcal{A}\right) \\
&\leqq \lim_k \frac{1}{k}F\left(\bigvee_{l=1}^{2n} T^l\mathcal{A}\right)+F\left(\bigvee_{l=-k+1}^{0} T^l\mathcal{A}\right) \\
&= t(T^{-1},\mathcal{A};F).
\end{aligned}
$$

Therefore

$$t(T^{-1}, F)=\underline{\lim}_n\, t(T^{-1}, \mathcal{B}_n; F)\leqq t(T^{-1}, \mathcal{A}; F)\leqq t(T^{-1}, F). \qquad \text{Q.E.D.}$$

PROPOSITION 6. 2. *If T is a Bernoulli shift*[3] *determined by a probability vector* $(p_1, \cdots, p_n)$, *then*

$$t(T^{-1}; F)=f_n(p_1, \cdots, p_n).$$

Proof. Let $\mathcal{A}_0\in\mathfrak{F}$ be a time 0 field.[4] Then $\mathcal{A}_0$ satisfies the assumption of Prop. 1, and by the condition V,

$$F(\mathcal{A}\vee T^{-1}\mathcal{A}\vee\cdots\vee T^{-k+1}\mathcal{A})=F(\mathcal{A})+F(T^{-1}\mathcal{A})+\cdots+F(T^{-k+1}\mathcal{A})$$
$$=kF(\mathcal{A}),$$

hence

$$t(T^{-1}; F)=t(T^{-1}, \mathcal{A}_0; F)=F(\mathcal{A}_0)=f_n(p_1, \cdots, p_n). \qquad \text{Q.E.D.}$$

Proof of theorem 1. We consider two Bernoulli shifts T_1 and T_2 determined by probability vectors $(p_1, \cdots, p_n)$ and $(q_1, \cdots, q_m)$ respectively. We express these Bernoulli schemes as the quartets $(\rho^I, \mathcal{X}, p, T_1)$ and $(\sigma^I, \mathcal{Y}, q, T_2)$.

If $-\sum_{i=1}^n p_i \log p_i=-\sum_{j=1}^m q_j \log q_j$, then by the Sinai's theorem [9], the Bernoulli shifts are weakly isomorphic, i.e., there exists two measure preserving transformations φ and ψ, which are from ρ^I to σ^I and from σ^I to ρ^I respectively, and

$$\varphi T_1=T_2\varphi \quad \text{and} \quad \psi T_2=T_1\psi.$$

We can regard φ^{-1} and ψ^{-1} as the semilattice morphisms from $\mathfrak{F}_2$ to $\mathfrak{F}_1$ and from $\mathfrak{F}_1$ to $\mathfrak{F}_2$ respectively, where $\mathfrak{F}_1$ is a class of all measurable finite partitions of ρ^I and $\mathfrak{F}_2$ is of σ^I. Then $(\mathfrak{F}_1, T_1^{-1})$ and $(\mathfrak{F}_2, T_2^{-1})$ are weakly isomorphic in the sense of AD-system by two semilattice morphisms φ^{-1} and ψ^{-1}, where the semivaluations on them are F-functions made by an arbitrarily fixed $\{f_n\}$ satisfying I$\sim$V.

By Prop. 1. 2, and Th. 5. 1,

$$f_n(p_1, \cdots, p_n)=t(T_1^{-1}; F)=t(T_2^{-1}; F)=f_m(q_1, \cdots, q_m).$$

Now to a positive real number λ we give a value $f_n(p_1, \cdots, p_n)$ if $\lambda=-\sum_{i=1}^n p_i \log p_i$. Then this correspondence $\lambda\to f_n(p_1, \cdots, p_n)$ is independent from the way of choice $(p_1 \cdots, p_n)$ and uniquely determined. We write the correspondence $g(\cdot)$, then the function $g(\cdot)$ is defined from non-negative real numbers R^+ to real numbers R. And we can see that $g(\cdot)$ has lineality;

$$g(\lambda_1+\lambda_2)=g(\lambda_1)+g(\lambda_2) \qquad (\lambda_1, \lambda_2\in R^+),$$

because for two probability vectors (q_j) and (r_k) with

3) The concept of the Bernoulli shift is explained in [1], p. 3 and p. 6.

4) Also explained in [1].

$$\lambda_1 = -\sum_j q_j \log q_j \qquad \text{and} \qquad \lambda_2 = -\sum_k r_k \log r_k,$$

$$\lambda_1 + \lambda_2 = -\sum_{j,k} q_j r_k \log q_j r_k,$$

hence

$$g(\lambda_1+\lambda_2) = f_{nm}(\{q_j r_k\}) = f_n(\{q_j\}) + f_m(\{r_k\}) = g(\lambda_1) + g(\lambda_2).$$

Next we show that the function $g(\cdot)$ is monotone increasing. Let us assume $\lambda_1 < \lambda_2$ ($\lambda_1, \lambda_2 \in R^+$), then there exists a vector $(p_1, \cdots, p_n)$ and $\lambda_1 = -\sum_{i=1}^n p_i \log p_i$. In this case we can find the number $\{x_k\}_{k=1}^m$ satisfying

$$\text{(6. 6)} \qquad \sum_{k=1}^m x_k = p_n, \qquad x_k \geqq 0 \quad (k=1, \cdots, m)$$

and $\lambda_2 = -\sum_{i=1}^{n-1} p_i \log p_i - \sum_{k=1}^m x_k \log x_k$; in fact, $\mu = (\lambda_2 - \lambda_1)/p_n$ is positive and there exists a probability vector $(p_1', \cdots, p_m')$ and

$$\mu = (\lambda_2 - \lambda_1)/p_n = -\sum_{k=1}^m p_k' \log p_k',$$

then writing $x_k = p_n p_k'$, we can see (6. 6) and

$$\begin{aligned} &-\sum_{i=1}^{n-1} p_i \log p_i - \sum_{k=1}^m (p_n p_k') \log (p_n p_k') \\ &= -\sum_{i=1}^{n-1} p_i \log p_i - p_n \log p_n - p_n \left(\sum_{k=1}^m p_k' \log p_k' \right) \\ &= \lambda_1 + p_n \left(\frac{\lambda_2 - \lambda_1}{p_n} \right) = \lambda_2. \end{aligned}$$

Now we consider two Bermoulli shifts determined by the probability vectors $(p_1, \cdots, p_n)$ and $(p_1, \cdots, p_{n-1}, x_1, \cdots, x_m)$ respectively. Then there exists a self-evident semilattice morphism from a semilattice of finite partitions of the former to the semilattice of the latter. Hence by Corollary of Th. 5. 1,

$$\begin{aligned} g(\lambda_2) &= f_{n+m-1}(p_1, \cdots, p_{n-1}, x_1, \cdots, x_m) \\ &\geqq f_n(p, \cdots, p_n) = g(\lambda_1). \end{aligned}$$

Then we can easily see that the function $g(\cdot)$ is written

$$g(\lambda) = c\lambda \qquad (c>0).$$

Therefore

$$f_n(p_1, \cdots, p_n) = -c \sum_{i=1}^n p_i \log p_i. \qquad \text{Q.E.D.}$$

BIBLIOGRAPHY

[1] BILLINGSLEY, P., Ergodic theory and information. John Wiley & Sons (1965).
[2] BIRKHOFF, G., Lattice theory. Amer. Math. Soc. Colloq. Publ. **25** (1948).
[3] FADDEEV, A. D., On the notion of entropy of a finite probability space. Uspehi Mat. Nauk. **11** (1956), 227–231. (in Russian)
[4] HINCHIN, A. I., Mathematical foundations of information theory. Dover Publ. (1958). (English translation)
[5] INGARDEN, R. S., Information theory and thermodynamic of light, I. Fortschritte der Physik **12** (1964), 567–594.
[6] RÉNYI, A., On measures of entropy and information. Proc. 4th Berkeley Symp. I (1962), 547-561.
[7] ROKHLIN, V. S., Generators in ergodic theorems. Vestnik Leningrad Univ. (1963), 26–32. (in Russian)
[8] SHANNON, C. E., A mathematical theory of communication. Bell system. Tech. Journ. **27** (1948), 370–423, 623–656.
[9] SINAI, JA. G., Weak isomorphism of transformations with invariant measure. Amer. Math. Soc. Transl. **2** (1966), 123–143. (English translation)
[10] TVERBERG, H., A new derivation of the information function. Math. Scand. **6** (1958), 297–298.

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.