# Mean Entropy of States in Classical Statistical Mechanics

Derek W. Robinson

CERN — Geneva

David Ruelle

I.H.E.S., Bures-sur-Yvette, France

**Abstract.** The equilibrium states for an infinite system of classical mechanics may be represented by states over Abelian $C^*$ algebras. We consider here continuous and lattice systems and define a mean entropy for their states. The properties of this mean entropy are investigated: linearity, upper semi-continuity, integral representations. In the lattice case, it is found that our mean entropy coincides with the Kolmogorov-Sinai invariant of ergodic theory.

## 0. Introduction

A new approach to the description of the equilibrium states of statistical mechanics has recently been intensively studied. In this approach these states are identified with states on a $B^*$ algebra $\mathfrak{A}$. It is assumed that the theory is invariant under a group $G$ (for instance the Euclidean or translation group) and the states considered are $G$ invariant. The algebra $\mathfrak{A}$ is Abelian for classical systems and non-Abelian for quantum systems. $G$ invariant states on Abelian $C^*$ algebras may be identified with measures on a compact set which are invariant under a group of homeomorphisms of this set, their study is thus naturally part of ergodic theory. Many of the recent results have consisted in extending ergodic theory to the case of a non-Abelian algebra $\mathfrak{A}$. It would thus be natural to obtain a non-Abelian extension of the mean entropy introduced by Kolmogorov and Sinaï (Kolmogorov-Sinaï invariant). Another reason for doing this is that a mean entropy should, on physical grounds, be associated with the equilibrium states of statistical mechanics (see [10]). In this paper we undertake the more modest project of giving a natural physical definition of mean entropy for classical systems, studying its properties and finding its relations with the Kolmogorov-Sinaï invariant[1].

## 1. States of Classical Statistical Mechanics

The description of equilibrium states in statistical mechanics as states on $B^*$ algebras has been considered recently by several authors

---

[1] Results similar to those described here have been obtained independently by O. Lanford (unpublished).

[3, 5, 6, 7, 10]. We summarize here briefly some facts pertaining to the case of classical statistical mechanics [11]. For simplicity we shall ignore the description of momenta of particles and assume that the one-particle configuration space $T$ is either $R^\nu$ (continuous systems) or $Z^\nu$ (lattice systems) where $R$ (the reals) and $Z$ (the integers) have the usual topologies. The invariance group of the theory is that of translations ($\equiv T$).

The Cartesian product of $n$ copies of $T$ is noted $T^n$ and the sum $\sum\limits_{n \geq 0} T^n$ of disjoint copies of all $T^n$ is noted $\mathscr{T}$. Let $\Lambda \subset T$ be a bounded open set (i. e., a finite set if $T = Z^\nu$). We call $\mathscr{K}^n_\Lambda$ the space of real continuous functions on $T^n$ with support in $\Lambda^n$, we call $\mathscr{K}_\Lambda$ the space of sequences $(f^n)_{n \geq 0}$ where $f^n \in \mathscr{K}^n_\Lambda$ and $f^n = 0$ for $n$ large enough, and we call $\mathscr{K}$ the union of the $\mathscr{K}_\Lambda$. An element of $\mathscr{K}$ may thus be considered as a function on $\mathscr{T}$.

For every bounded open $\Lambda \subset T$, and integer $n \geq 0$, let $\mu^n_\Lambda \geq 0$ be a measure on $\Lambda^n$, symmetric in its $n$ arguments. We shall say that $(\mu^n_\Lambda)$ is a *family of density distributions* if the following conditions are satisfied.

(D 1) **Normalization.** For all $\Lambda$

$$\sum_{n=0}^{\infty} \mu^n_\Lambda(\Lambda^n) = 1 \ . \tag{1.1}$$

(D 2) **Compatibility.** Let $\Lambda \subset \Lambda'$ and $\chi_{\Lambda'/\Lambda}$ be the characteristic function of $\Lambda'/\Lambda$ where $\Lambda'/\Lambda$ is the complement of $\Lambda$ in $\Lambda'$. If $f^n \in \mathscr{K}^n_\Lambda$, then

$$\mu^n_\Lambda(f^n) = \sum_{m=0}^{\infty} \frac{(n+m)!}{n!\,m!} \mu^{n+m}_{\Lambda'}(f^n \otimes \chi^{\otimes m}_{\Lambda'/\Lambda}) \tag{1.2}$$

where

$$[f^n \otimes \chi^{\otimes m}_{\Lambda'/\Lambda}] \, (x_1, \ldots, x_{n+m}) = f^n(x_1, \ldots, x_n) \, \chi_{\Lambda'/\Lambda}(x_{n+1}) \cdots \chi_{\Lambda'/\Lambda}(x_{n+m}) \ .$$

Notice that $\Lambda^0$ is reduced to a point even if $\Lambda = \phi$ is the empty set and, as $\phi^n = \phi$ for $n > 0$, (D 1) gives

$$\mu^0_\phi(\phi^0) = 1 \ . \tag{1.3}$$

Inserting this formula in (D 2) yields again (D 1).

If $\mathscr{K} \ni f = (f^n)_{n \geq 0}$, a function $Sf$ on $\mathscr{T}$ may be defined so that its restriction to $T^n$ is

$$Sf(x_1, \ldots, x_n) = \sum_{p \geq 0} \sum_{i_1 = 1}^{n} \cdots \sum_{i_p = 1}^{n} f^p(x_{i_1}, \ldots, x_{i_p}) \ .$$

For any integer $q \geq 0$; $f_1, \ldots, f_q \in \mathscr{K}$ and bounded continuous complex function $\varphi$ on $R^q$, consider the function $\varphi(Sf_1, \ldots, Sf_q)$ on $\mathscr{T}$. With respect to the usual operations on functions and the * operation given by complex conjugation, such functions form a commutative * algebra $\widetilde{\mathfrak{A}}$. The closure $\mathfrak{A}$ of $\widetilde{\mathfrak{A}}$ with respect to the uniform norm is an Abelian $B^*$

algebra. Given a family of density distributions $(\mu_\Lambda^n)$, a state $\varrho$ on $\mathfrak{A}$ is defined as follows. For each $\varphi(Sf_1, \ldots, Sf_q)$ let $\Lambda$ be such that $f_1, \ldots, f_q \in \mathscr{K}_\Lambda$, then let

$$\varrho(\varphi(Sf_1, \ldots, Sf_q))$$
$$= \sum_{n \geq 0} \int d\mu_\Lambda^n(x_1, \ldots, x_n) \, \varphi(Sf_1(x_1, \ldots, x_n), \ldots, Sf_q(x_1, \ldots, x_n)) \, .$$

It can be checked that this definition is independent of the choice of $\Lambda$ and extends by continuity to a state on $\mathfrak{A}$.

The mapping $(\mu_\Lambda^n) \to \varrho$ is injective, we call $\mathscr{F}$ its image in the set $E$ of all states. The states $\varrho$ in $\mathscr{F}$ are characterized by the fact that for every $\varepsilon > 0$ and $f \in \mathscr{K}$ one can find a continuous function $\varphi$ on $R$ with values in $[0, 1]$ and compact support such that

$$\varrho(\varphi(Sf)) > 1 - \varepsilon \, .$$

Given a function $F$ on $\mathscr{T}$, a translation $\tau_a$ by $a \in T$ is defined by

$$\tau_a F(x_1, \ldots, x_n) = F(x_1 - a, \ldots, x_n - a) \, .$$

In particular the $\tau_a$ yields a group of automorphisms of $\mathfrak{A}$. We call $\mathscr{L}^\perp$ the subspace of the dual $\mathfrak{A}'$ of $\mathfrak{A}$ consisting of the invariant forms $f : f(\tau_a A) = f(A)$. The set $E \cap \mathscr{L}^\perp$ of invariant states is compact for the $w^*$ topology of $\mathfrak{A}'$. The set $\mathscr{F} \cap \mathscr{L}^\perp$ consists of the images of the families of those density distributions which satisfy the requirement

(D 3) **Invariance.** If $a \in T$ and $f^n \in \mathscr{K}_\Lambda^n$, then

$$\mu_\Lambda^n(f^n) = \mu_{\Lambda+a}^n(\tau_a f^n) \, . \tag{1.4}$$

## 2. Entropy for Continuous Systems

In this Section we take $T = R^\nu$ (continuous system). Let $(\mu_\Lambda^n)$ be a family of density distributions and assume that for every $\Lambda$, $n$ the measure $\mu_\Lambda^n$ is absolutely continuous with respect to the Lebesgue measure. If $V(\Lambda)$ is the Lebesgue measure (volume) of $\Lambda$, we write

$$d\mu_\Lambda^n(x_1, \ldots, x_n) = \frac{e^{-V(\Lambda)}}{n!} f_\Lambda^n(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \, . \tag{2.1}$$

We shall also write $f_\Lambda = (f_\Lambda^n)_{n \geq 0}$ and use the notation

$$\int d_\Lambda x \cdot = \sum_{n \geq 0} \frac{e^{-V(\Lambda)}}{n!} \int_{\Lambda^n} dx_1 \ldots dx_n \, . \tag{2.2}$$

Then

$$\int d_\Lambda x = 1 \tag{2.3}$$

and (D 1) becomes

$$\int d_\Lambda x \, f_\Lambda(x) = 1 \, . \tag{2.4}$$

Let $\Lambda' \supset \Lambda$. If $x^{(1)} \in \sum_{n_1 \geq 0} \Lambda^{n_1}$ and $x^{(2)} \in \sum_{n_2 \geq 0} (\Lambda'/\Lambda)^{n_2}$, we can identify $(x^{(1)}, x^{(2)})$ with a point of $\sum_{n \geq 0} (\Lambda')^n$ by $((x_1^{(1)}, \ldots, x_{n_1}^{(1)}), (x_1^{(2)}, \ldots, x_{n_2}^{(2)})) \to$ $\to (x_1^{(1)}, \ldots, x_{n_1}^{(1)}, x_1^{(2)}, \ldots, x_{n_2}^{(2)})$. We define $d_{\Lambda'/\Lambda} x^{(2)}$ by analogy with (2.2), using then the symmetry of the $f_\Lambda^n$, the compatibility condition (D 2) becomes

$$f_\Lambda(x^{(1)}) = \int d_{\Lambda'/\Lambda} x^{(2)} f_{\Lambda'}(x^{(1)}, x^{(2)}) . \tag{2.5}$$

We define now an entropy $S(\Lambda)$ by

$$S(\Lambda) = - \int d_\Lambda x \, f_\Lambda(x) \log f_\Lambda(x) . \tag{2.6}$$

Notice that we may have $S(\Lambda) = -\infty$ and that (1.3) yields

$$S(\phi) = 0 . \tag{2.7}$$

**Proposition 1.** *The following inequalities hold*

*Negativity:* $S(\Lambda) \leq 0$;

*Decrease:* $\Lambda' \supset \Lambda \Rightarrow S(\Lambda') - S(\Lambda) \leq 0$;

*Strong sub-additivity:* $S(\Lambda \cup \Lambda') - S(\Lambda) - S(\Lambda') + S(\Lambda \cap \Lambda') \leq 0$.

The convexity of the function $t \to t \log t \ (t > 0)$ implies

$$-t \log t \leq 1 - t$$

and hence if $\Lambda' \subset \Lambda$ we obtain, with the help of (2.5)

$$S(\Lambda') - S(\Lambda) = - \int d_\Lambda x^{(1)} \int d_{\Lambda'/\Lambda} x^{(2)} f_{\Lambda'}(x^{(1)}, x^{(2)}) \log \left[ \frac{f_{\Lambda'}(x^{(1)}, x^{(2)})}{f_\Lambda(x^{(1)})} \right] \leq$$

$$\leq \int\int f_{\Lambda'} \left[ \frac{f_\Lambda}{f_{\Lambda'}} - 1 \right] = 0$$

where we have restricted the integrations to the region where $f_{\Lambda'} < 0$. This proves the decrease property of $S(\Lambda)$ and choosing $\Lambda = \phi$, also the negativity.

To prove strong sub-additivity we use variables $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ corresponding respectively to $\Lambda \cap \Lambda'$, $\Lambda/\Lambda' \ (\equiv \Lambda/(\Lambda' \cap \Lambda))$, $\Lambda'/\Lambda \ (\equiv \Lambda'/(\Lambda \cap \Lambda'))$, then

$$S(\Lambda \cup \Lambda') - S(\Lambda) - S(\Lambda') + S(\Lambda \cap \Lambda')$$

$$= - \int d_{\Lambda \cap \Lambda'} x^{(1)} \int d_{\Lambda'/\Lambda} x^{(2)} \int d_{\Lambda/\Lambda'} x^{(3)} f_{\Lambda \cup \Lambda'}(x^{(1)}, x^{(2)}, x^{(3)}) \times$$

$$\times \log \left[ \frac{f_{\Lambda' \cup \Lambda}(x^{(1)}, x^{(2)}, x^{(3)}) f_{\Lambda \cap \Lambda'}(x^{(1)})}{f_\Lambda(x^{(1)}, x^{(2)}) f_{\Lambda'}(x^{(1)}, x^{(3)})} \right] \leq$$

$$\leq \int\int\int \left[ \frac{f_\Lambda f_{\Lambda'}}{f_{\Lambda \cap \Lambda'}} - f_{\Lambda \cup \Lambda'} \right] = \int\int \frac{f_\Lambda}{f_{\Lambda \cap \Lambda'}} \times$$

$$\times \int d_{\Lambda'/\Lambda} x^{(3)} f_{\Lambda'}(x^{(1)}, x^{(3)}) - \int\int\int f_{\Lambda \cup \Lambda'} = \int\int f_\Lambda - \int\int\int f_{\Lambda \cup \Lambda'} = 0$$

where we have restricted the integrations to the region where $f_{\Lambda \cup \Lambda'} > 0$.

*Remark.* If $\Lambda$ and $\Lambda'$ are disjoint we have by (2.7)

$$S(\Lambda \cup \Lambda') \leqq S(\Lambda) + S(\Lambda')$$

i. e., *sub-additivity*. Notice also that if $\Lambda$ and $\Lambda'$ differ by a set of Lebesgue measure zero, then $S(\Lambda) = S(\Lambda')$.

If $a = (a^1, \ldots, a^\nu) \in T$ and $a^1 \geqq 0, \ldots, a^\nu \geqq 0$ we let

$$\Lambda(a) = \{x \in T : 0 < x^i < a^i \quad \text{for} \quad i = 1, \ldots, \nu\} \,.$$

The translates of $\Lambda(a)$ by vectors $(n^1 a^1, \ldots, n^\nu a^\nu)$ where $(n^1, \ldots, n^\nu) \in Z^\nu$ form a partition $\mathscr{P}_a$ of $T$ (up to sets of Lebesgue measure zero). Let $n_\Lambda^+(a)$ (resp. $n_\Lambda^-(a)$) be the number of sets of this partition which have non-void intersection with $\Lambda$ (resp. which are contained in $\Lambda$) and let $\Gamma_\Lambda^+(a)$ (resp. $\Gamma_\Lambda^-(a)$) be the union of these sets.

**Definition 1.** *We say that the (bounded open) sets $\Lambda$ tend to infinity in the sense of* VAN HOVE *and we write $\Lambda \to \infty$ if for every partition $\mathscr{P}_a$*

$$\lim \frac{n_\Lambda^+(a)}{n_\Lambda^-(a)} = 1 \,.$$

**Proposition 2.** *If the family $(\mu_\Lambda^n)$ satisfies the invariance condition* (D 3), *then*

$$s = \lim_{\Lambda \to \infty} \frac{S(\Lambda)}{V(\Lambda)} \tag{2.8}$$

*exists, $s \in [-\infty, 0]$.*

By the decrease and sub-additivity properties of $S$ we have

$$\frac{S(\Lambda)}{V(\Lambda)} \leqq \frac{S(\Gamma_\Lambda^-(a))}{V(\Gamma_\Lambda^+(a))} \leqq \frac{n_\Lambda^-(a)}{n_\Lambda^+(a)} \frac{S(\Lambda(a))}{V(\Lambda(a))} \,. \tag{2.9}$$

We define

$$s = \inf_a \frac{S(\Lambda(a))}{V(\Lambda(a))} \,. \tag{2.10}$$

In particular if $s = -\infty$, (2.8) follows from (2.9). Let thus $s$ be finite, given $\varepsilon > 0$ we can choose $a_0$ such that

$$\frac{S(\Lambda(a_0))}{V(\Lambda(a_0))} < s + \varepsilon$$

and (2.9) yields

$$\frac{S(\Lambda)}{V(\Lambda)} \leqq \frac{n_\Lambda^-(a_0)}{n_\Lambda^+(a_0)} (s + \varepsilon) \,. \tag{2.11}$$

We construct now $\Gamma_\Lambda^+(a_0)$ by successively adding translates of $\Lambda(a_0)$ in the lexicographic order of the vectors $(n^1, \ldots, n^\nu)$ defining these translates. Let $\Gamma_n$ be the union of the first $n$ translates, so that $\Gamma_{n^+} = \Gamma_\Lambda^+(a_0)$. Let $b = (m^1 a_0^1, \ldots, m^\nu a_0^\nu)$ where $m^1, \ldots, m^\nu$ are positive integers, $\Lambda(b)$ is also a union of translates of $\Lambda(a_0)$ and can be constructed by adding them successively in lexicographic order, let here $\Delta_n$ be the union of the first $n$ translates.

If we assume that

$$S(\Gamma_{n+1}) - S(\Gamma_n) < (s - \varepsilon) \, V(\Lambda(a_0)) \tag{2.12}$$

the strong sub-additivity of $S$ implies that

$$S(\Delta_{N+1}) - S(\Delta_N) < (s - \varepsilon)\, V(\Lambda(a_0)) \qquad (2.13)$$

for all $N$ such that there exists a translation mapping $\Gamma_{n+1}$ into $\Delta_{N+1}$ with the property that the last translate of $\Lambda(a_0)$ in $\Gamma_{n+1}$ is mapped onto the last translate of $\Lambda(a_0)$ in $\Delta_{N+1}$. If $b$ is large enough, (2.13) will hold for almost all $N$ and, using the decrease property of $s$, we obtain

$$S(\Lambda(b)) = \sum_N \left( S(\Delta_{N+1}) - S(\Delta_N) \right) < \left( s - \frac{\varepsilon}{2} \right) V(\Lambda(b)) \qquad (2.14)$$

in contradiction with (2.10), therefore

$$S(\Gamma_{n+1}) - S(\Gamma_n) \geqq s\, V(\Lambda(a_0)) \,.$$

Hence, summing over $n$ we find

$$S(\Gamma_\Lambda^+(a_0)) \geqq n_\Lambda^+(a_0)\, s\, V(\Lambda(a_0))$$

and thus

$$\frac{S(\Lambda)}{V(\Lambda)} \geqq \frac{S(\Gamma_\Lambda^+(a_0))}{V(\Gamma_\Lambda^-(a_0))} \geqq \frac{n_\Lambda^+(a_0)}{n_\Lambda^-(a_0)}\, s \,. \qquad (2.15)$$

Comparison of (2.11) and (2.15) proves (2.8).

**Definition 2.** *A mean entropy* $s(\varrho) \in [-\infty, 0]$ *is defined for every* $\varrho \in E \cap \mathscr{L}^\perp$ *by*

    i) $s(\varrho) = -\infty$ *if* $\varrho \notin \mathscr{F} \cap \mathscr{L}^\perp$;

    ii) $s(\varrho) = -\infty$ *if* $\varrho \in \mathscr{F} \cap \mathscr{L}^\perp$ *but the measures* $\mu_\Lambda^n$ *associated with* $\varrho$ *are not all absolutely continuous with respect to the Lebesgue measure;*

    iii) $s(\varrho) = s$ *as defined in Proposition 2 otherwise.*

## 3. Properties of the Mean Entropy of Continuous Systems

**Proposition 3.** *The functional* $s(\cdot)$ *is affine on* $E \cap \mathscr{L}^\perp$.

Let $0 < \alpha < 1$ and $\varrho_1, \varrho_2 \in E \cap \mathscr{L}^\perp$. From the characterization of $\mathscr{F}$ in Section 1 it follows that if $\varrho_1$ or $\varrho_2$ falls under i) or ii) in Definition 2, then $\alpha \varrho_1 + (1 - \alpha) \varrho_2$ also falls under i) or ii). We may thus assume that $s(\varrho_1)$ and $s(\varrho_2)$ are defined by Proposition 2. We have then by the convexity of $t \log t$ and the increase of $\log t$

$$- \int [\alpha f_{1\Lambda} \log f_{1\Lambda} + (1 - \alpha) f_{2\Lambda} \log f_{2\Lambda}] \leqq$$

$$\leqq - \int (\alpha f_{1\Lambda} + (1 - \alpha) f_{2\Lambda}) \log(\alpha f_{1\Lambda} + (1 - \alpha) f_{2\Lambda}) \leqq$$

$$\leqq - \int [\alpha f_{1\Lambda} \log \alpha f_{1\Lambda} + (1 - \alpha) f_{2\Lambda} \log(1 - \alpha) f_{2\Lambda}] \qquad (3.1)$$

$$= - \int [\alpha f_{1\Lambda} \log f_{1\Lambda} + (1 - \alpha) f_{2\Lambda} \log f_{2\Lambda}] - \alpha \log \alpha - (1 - \alpha) \log(1 - \alpha) \leqq$$

$$\leqq - \int [\alpha f_{1\Lambda} \log f_{1\Lambda} + [(1 - \alpha) f_{2\Lambda} \log f_{2\Lambda}] + \log 2 \,.$$

Dividing by $V(\Lambda)$ and taking the limit $\Lambda \to \infty$ yields

$$s(\alpha \varrho_1 + (1 - \alpha) \varrho_2) = \alpha s(\varrho_1) + (1 - \alpha) s(\varrho_2) \,.$$

**Proposition 4.** *The functional* $s(\cdot)$ *is the lower bound of a family of concave continuous functionals on* $E \cap \mathscr{L}^{\perp}$ *equipped with the* $w^*$ *topology.*

Let $(\Psi_{\alpha})$ be a finite continuous partition of the unity on $\sum\limits_{n \geq 0} \Lambda^n$: $\Psi_{\alpha} \geqq 0$, $\sum \Psi_{\alpha} = 1$, then

$$- \int f_{\Lambda} \log f_{\Lambda} \leqq - \sum_{\alpha} \left( \int f_{\Lambda} \Psi_{\alpha} \right) \log \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} .$$

We have indeed by the convexity of $t \log t$

$$- \sum_{\alpha} \left( \int \Psi_{\alpha} \right) \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} \log \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} \geqq$$

$$\geqq - \sum_{\alpha} \left( \int \Psi_{\alpha} \right) \frac{\int d_{\Lambda} x \, \Psi_{\alpha}(x) \, f_{\Lambda}(x) \log f_{\Lambda}(x)}{\int \Psi_{\alpha}}$$

$$= - \int d_{\Lambda} x \left[ \sum_{\alpha} \Psi_{\alpha}(x) \right] f_{\Lambda}(x) \log f_{\Lambda}(x) .$$

Let now $\Psi_{\alpha} = \varphi_{\alpha}(S f_1, \ldots, S f_q)$ with $f_1, \ldots, f_q \in \mathscr{K}_{\Lambda}$, then $\int f_{\Lambda} \Psi_{\alpha} = \varrho(\varphi_{\alpha}(S f_1, \ldots, S f_q))$ and it is clear that for all $\varrho \in E \cap \mathscr{L}^{\perp}$

$$S(\varrho) \leqq \Phi(\varrho) = - V(\Lambda)^{-1} \sum_{\alpha} \varrho(\varphi_{\alpha}(S f_1, \ldots, S f_q)) \log \frac{\varrho(\varphi_{\alpha}(S f_1, \ldots, S f_q))}{\int \varphi_{\alpha}(S f_1, \ldots, S f_q)} \quad (3.2)$$

where $\Phi$ is concave and continuous. We show now that for every $\varrho \in E \cap \mathscr{L}^{\perp}$ we can choose $\Phi$ such that $\Phi(\varrho)$ is arbitrary close to $s(\varrho)$.

i) If $\varrho \notin \mathscr{F} \cap \mathscr{L}^{\perp}$ there exist $\delta > 0$ and $f \in \mathscr{K}_{\Lambda}$ such that for any continuous function $\varphi$ on $R$ with values in $[0, 1]$ and compact support

$$\varrho(\varphi(S f)) < 1 - \delta .$$

We choose a sequence $\varphi_1^{(n)}$ of such functions such that if $\varphi_2^{(n)} = 1 - \varphi_1^{(n)}$ we have

$$\int d_{\Lambda} x \, \varphi_2^{(n)}(S f(x)) \to 0 \quad \varrho(\varphi_2^n(S f)) \to \delta_1 > 0$$

then

$$\Phi^{(n)}(\varrho) = V(\Lambda)^{-1} \sum_{\alpha = 1}^{2} \varrho(\varphi_{\alpha}^{(n)}(S f)) \log \frac{\varrho(\varphi_{\alpha}^{(n)}(S f))}{\int \varphi_{\alpha}^{(n)}(S f)} \leqq$$

$$\leqq V(\Lambda)^{-1} \left[ 2 e^{-1} + \varrho(\varphi_1^{(n)}(S f)) \log \int \varphi_1^{(n)}(S f) + \right.$$

$$\left. + \varrho(\varphi_2^{(n)}(S f)) \log \int \varphi_2^{(n)}(S f) \right] \to - \infty .$$

ii) If $\varrho \in \mathscr{F} \cap \mathscr{L}^{\perp}$ but $\mu_{\Lambda}^n$ is not absolutely continuous with respect to the Lebesgue measure, the existence of a set with zero Lebesgue measure and non-vanishing $\mu_{\Lambda}$ measure implies the existence of a sequence $\varphi_2^{(n)}(S f^{(n)})$ such that

$$\int d_{\Lambda} x \, \varphi_2^{(n)}(S f^{(n)}(x)) \to 0 \quad \varrho(\varphi_2^{(n)}(S f^{(n)})) \to \delta_1 > 0$$

and the argument proceeds as in ii).

iii) If $\varrho \in \mathscr{F} \cap \mathscr{L}^{\perp}$ and the measures $\mu_{\Lambda}^{n}$ are absolutely continuous with respect to the Lebesgue measure, let again $f_{\Lambda}$ be the corresponding $L^1$ functions. We take $\Lambda$ such that $S(\Lambda)/V(\Lambda)$ is close to $s(\varrho)$. We can choose the partition of the unity $(\Psi_{\alpha})$ such that the function

$$\tilde{f}_{\Lambda}(\cdot) = \sum_{\alpha} \Psi_{\alpha}(\cdot) \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}}$$

approaches $f_{\Lambda}$ in the $L^1$ norm and we may assume that the $\Psi_{\alpha}$ are of the form $\varphi_{\alpha}(S f_1, \ldots, S f_q)$ with $f_1, \ldots, f_q \in \mathscr{K}_{\Lambda}$.

If $-\int f_{\Lambda} \log f_{\Lambda} = -\infty$ it is already clear that $\int \tilde{f}_{\Lambda} \log \tilde{f}_{\Lambda}$ will approach $\int f_{\Lambda} \log f_{\Lambda}$. We show now that this remains true if $\int f_{\Lambda} \log f_{\Lambda}$ is finite.

Let $\lambda > e$ and $\varphi_{\lambda}$ be the function defined on the real line by

$$\varphi_{\lambda}(t) = \begin{cases} \lambda^{-1} & \text{if} \quad t \leq \lambda^{-1} \\ t & \text{if} \quad \lambda^{-1} \leq t \leq \lambda \\ \lambda & \text{if} \quad \lambda \leq t. \end{cases}$$

Given $\varepsilon > 0$ we may choose $\lambda$ such that

$$\int [f_{\Lambda} \log f_{\Lambda} - \varphi_{\lambda}(f_{\Lambda}) \log \varphi_{\lambda}(f_{\Lambda})] < \varepsilon/4 \qquad (3.3)$$

and because of the inequalities

$$1 + \log \lambda < 2 |1 + \log \lambda^{-1}| \qquad (3.4)$$

$$t \log t - t_0 \log t_0 \geq (t - t_0)(1 + \log t_0) \qquad (3.5)$$

we get

$$(1 + \log \lambda) \int |f_{\Lambda} - \varphi_{\lambda}(f_{\Lambda})| \leq 2 \int (f_{\Lambda} - \varphi_{\lambda}(f_{\Lambda}))(1 + \log \varphi_{\lambda}(f_{\Lambda})) < \varepsilon/2 . \qquad (3.6)$$

Using (3.3), (3.4), (3.5), (3.6) we obtain

$$\int (f_{\Lambda} \log f_{\Lambda} - \tilde{f}_{\Lambda} \log \tilde{f}_{\Lambda})$$
$$= \int (f_{\Lambda} \log f_{\Lambda} - \varphi_{\lambda}(f_{\Lambda}) \log \varphi_{\lambda}(f_{\Lambda})) + \int (\varphi_{\lambda}(f_{\Lambda}) \log \varphi_{\lambda}(f_{\Lambda}) - \tilde{f}_{\Lambda} \log \tilde{f}_{\Lambda}) <$$
$$< \varepsilon/4 + \int (\varphi_{\lambda}(f_{\Lambda}) - \tilde{f}_{\Lambda})(1 + \log \varphi_{\lambda}(f_{\Lambda})) \leq$$
$$\leq \varepsilon/4 + (1 + \log \lambda) [\int |\varphi_{\lambda}(f_{\Lambda}) - f_{\Lambda}| + \int |f_{\Lambda} - \tilde{f}_{\Lambda}|] < \varepsilon$$

where we have chosen the $\Psi_{\alpha}$ such that

$$(1 + \log \lambda) \int |f_{\Lambda} - \tilde{f}_{\Lambda}| < \varepsilon/4 .$$

The proof is concluded by the remark that $-V(\Lambda)^{-1} \int \tilde{f}_{\Lambda} \log \tilde{f}_{\Lambda} \geq \Phi(\varrho)$, so that $\Phi(\varrho)$ is close to $S(\Lambda)/V(\Lambda)$ hence to $s(\varrho)$. By the convexity of $t \log t$ we have indeed

$$-V(\Lambda)^{-1} \int \tilde{f}_{\Lambda} \log \tilde{f}_{\Lambda}$$
$$= -V(\Lambda)^{-1} \int d_{\Lambda} x \left( \sum_{\alpha} \Psi_{\alpha}(x) \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} \right) \log \left( \sum_{\alpha} \Psi_{\alpha}(x) \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} \right) \geq$$
$$\geq -V(\Lambda)^{-1} \int d_{\Lambda} x \sum_{\alpha} \Psi_{\alpha}(x) \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} \log \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}}$$
$$= -V(\Lambda)^{-1} \sum_{\alpha} \left( \int f_{\Lambda} \Psi_{\alpha} \right) \log \frac{\int f_{\Lambda} \Psi_{\alpha}}{\int \Psi_{\alpha}} = \Phi(\varrho) .$$

**Proposition 5.** *The functional $s$ is affine upper semi-continuous on $E \cap \mathscr{L}^\perp$. If $\mu_\varrho$ is a measure on $E \cap \mathscr{L}^\perp$ with resultant $\varrho$, then $s(\varrho) = \mu_\varrho(s)$.*

$s$ is affine by Proposition 3, upper semi-continuous as the lower bound of a family of continuous functions by Proposition 4. That $s(\varrho) = \mu_\varrho(s)$ results from the proof of Lemma 10 in [2].

The formula $s(\varrho) = \mu_\varrho(s)$ is especially interesting when $\mu_\varrho$ is the (unique) maximal measure with resultant $\varrho$. In particular, it is known (see [11]) that if $\varrho \in \mathscr{F} \cap \mathscr{L}^\perp$, then $\mu_\varrho$ is carried by the set $\mathscr{F} \cap \mathscr{E} (E \cap \cap \mathscr{L}^\perp)$ where $\mathscr{E} (E \cap \mathscr{L}^\perp)$ is the set of extremal invariant states (or ergodic states [12]). Therefore if $\varrho \in \mathscr{F} \cap \mathscr{L}^\perp$, $s(\varrho)$ has an integral representation on the set of extremal invariant states.

If further it is possible to obtain the equilibrium states as solutions of a variational problem involving the entropy, the following points may be important.

1. An upper semi-continuous function defined over a compact set reaches its maximum.

2. An affine upper semi-continuous function defined over a convex compact set reaches its maximum at an extremal point (corresponding to a single thermodynamic phase).

## 4. Lattice Systems

In this Section we take $T = Z^\nu$ (lattice system). A bounded open set $\Lambda$ is now simply a finite subset of $T$: $\Lambda = \{x_1, \ldots, x_V\}$. Consider a point $(x_{i_1}, \ldots, x_{i_n})$ of $\Lambda^n$ and let $n_1, \ldots, n_V$ be the number of indices $i_k$ equal to $1, \ldots, V$. The measure $\mu_\Lambda^n$ is determined by the numbers $\mu_\Lambda^n(\{(x_{i_1}, \ldots, x_{i_n})\})$ and, since $\mu_\Lambda^n$ is symmetric in its arguments, $\mu_\Lambda^n(\{(x_{i_1}, \ldots, x_{i_n})\})$ depends only on $n_1, \ldots, n_V$. We write

$$f_\Lambda(n_1, \ldots, n_V) = \frac{n!}{n_1! \ldots n_V!} \, \mu_\Lambda^n(\{(x_{i_1}, \ldots, x_{i_n})\}) \,. \tag{4.1}$$

Notice that in this formula $\sum_{i=1}^V n_i = n$ and that $n!/(n_1! \ldots n_V!)$ is the number of points of $\Lambda^n$ which correspond to the same $n_1, \ldots, n_V$.

With this notation (D 1) becomes

$$\sum_{n_1=0}^\infty \cdots \sum_{n_V=1}^\infty f_\Lambda(n_1, \ldots, n_V) = 1 \,. \tag{4.2}$$

If $\Lambda' \supset \Lambda$ and $\Lambda'$ contains $V'$ points, (D 2) becomes

$$f_\Lambda(n_1, \ldots, n_V) = \sum_{n_{V+1}=0}^\infty \cdots \sum_{n_{V+V'}=0}^\infty f_{\Lambda'}(n_1, \ldots, n_V, n_{V+1}, \ldots, n_{V+V'}). \tag{4.3}$$

With $x = (n_1, \ldots, n_V)$ and

$$\int d_\Lambda \, x \cdot = \sum_{n_1=0}^\infty \cdots \sum_{n_V=0}^\infty \cdot \tag{4.4}$$

we may rewrite (4.2) and (4.3) as (2.4) and (2.5). There is however an important difference between the lattice and the continuous case, namely that whilst $d_A x$ is normalized in the continuous case by (2.3), it is in the lattice case a measure on a discrete space giving a mass 1 to each point.

We define now an entropy $S(A)$ by

$$S(A) = -\int d_A x\, f_A(x) \log f_A(x) = -\sum_{n_1 = 0}^{\infty} \cdots \sum_{n_V = 0}^{\infty} f_A(n_1, \ldots, n_V) \times \qquad (4.5)$$
$$\times \log f_A(n_1, \ldots, n_V) .$$

Notice that we may have $S(A) = +\infty$ and that (1.3) yields again

$$S(\phi) = 0 . \qquad (4.6)$$

**Proposition 6.** *The following inequalities hold*
*Positivity:* $S(A) \geqq 0$;
*Increase:* $A' \supset A \Rightarrow S(A') - S(A) \geqq 0$;
*Strong sub-additivity:* $S(A \cup A') - S(A) - S(A') + S(A \cap A') \leqq 0$.
From (4.2) one gets $0 \leqq f_A(n_1, \ldots, n_V) \leqq 1$, hence

$$-f_A(n_1, \ldots, n_V) \log f_A(n_1, \ldots, n_V) \geqq 0$$

which proves the positivity of $S$.

For the increase let $A' \supset A$, then we obtain, using (4.3)

$$f_{A'}(x^{(1)}, x^{(2)}) \leqq f_A(x^{(1)})$$

hence

$$S(A') - S(A) = -\int d_A x^{(1)} \int d_{A'/A} x^{(2)} f_{A'}(x^{(1)}, x^{(2)}) \left[\log f_{A'}(x^{(1)}, x^{(2)}) - \right.$$
$$\left. - \log f_A(x^{(1)})\right] \geqq 0 .$$

The proof of strong sub-additivity of the continuous case holds again here because it makes no use of the normalization of $d_A x$.

If $a = (a^1, \ldots, a^\nu) \in T$ and $a^1 \geqq 0, \ldots, a^\nu \geqq 0$ we let

$$A(a) = \{x \in T : 0 \leqq x^i < a^i \quad \text{for} \quad i = 1, \ldots, \nu\} .$$

Convergence in the sense of Van Hove can be defined as in the continuous case and we have

**Proposition 7.** *If the family $(\mu_A^n)$ satisfies the invariance condition (D 3), then*

$$s = \lim_{A \to \infty} \frac{S(A)}{V(A)}$$

*exists,* $s \in [0, +\infty]$.

We do not reproduce the proof which is analogous to that of Proposition 2, differing from it essentially only by the interchange of the superscripts $\pm$ in the formulas. We note however that to obtain (2.14) we use instead of the decrease of $S$ the inequality

$$S(A_{N+1}) - S(A_N) \leqq S(A(a_0)) < V(A(a_0))(s + \varepsilon) .$$

*Remark.* If, for some $\Lambda \neq \phi$, $S(\Lambda) < +\infty$ then, by the increase property of $S$, $S(\Lambda(1)) < +\infty$, where $1 = (1, \ldots, 1)$, then

$$s = \inf_a \frac{S(\Lambda(a))}{V(\Lambda(a))} \leqq \frac{S(\Lambda(1))}{V(\Lambda(1))} < +\infty .$$

**Definition 3.** *A mean entropy* $s(\varrho) \in [0, +\infty]$ *is defined for all* $\varrho \in \in \mathscr{F} \cap \mathscr{L}^\perp$ *by Proposition 7.*

**Proposition 8.** *The functional* $s(\cdot)$ *is affine on* $\mathscr{F} \cap \mathscr{L}^\perp$.

The proof is again given by (3.1).

Now by the Gel'fand isomorphism, any $\varrho \in E$ is identifiable with a measure $m_\varrho$ on the spectrum $\mathscr{E}(E)$ of $\mathfrak{A}$. If $\varrho \in \mathscr{F}$, then $m_\varrho$ is carried by $\mathscr{F} \cap \mathscr{E}(E)$ (see [10], Section 11) and conversely, $\mathscr{F} \cap \mathscr{E}(E)$ is a $G_\delta$ (countable intersection of open sets). If $\varrho \in \mathscr{F} \cap \mathscr{L}^\perp$, then the measure $m_\varrho$ on $\mathscr{F} \cap \mathscr{E}(E)$ is invariant under the transformations $\tau'_x$ of $\mathscr{F} \cap \mathscr{E}(E)$ associated with the automorphisms $\tau_x (x \in T)$ of $\mathfrak{A}$. Let $B$ be the $\sigma$ field on $\mathscr{F} \cap \mathscr{E}(E)$ induced by the $\sigma$ field of Baire sets on $\mathscr{E}(E)$ associated with the $w^*$ topology.

The quadruple $(\mathscr{F} \cap \mathscr{E}(E), B, m_\varrho, T))$ is a dynamical system in the sense of ergodic theory (see [4], Section 10), it is therefore natural to consider the concept of mean entropy introduced by Kolmogorov and Sinaï in this framework. For details of this theory we refer the reader to Jacobs [4], Billingsley [1], and Rokhlin [9] and papers quoted therein. The theory of the Kolmogorov-Sinaï invariant is usually developed for a group $T = Z$ (or $T = R$), but many of the results extend to $T = Z^\nu$ (see [8]) and will be used without further discussion.

**Proposition 9.** *If* $\varrho \in \mathscr{F} \cap \mathscr{L}^\perp$ *and* $s(\varrho) < +\infty$, *then the mean entropy* $s(\varrho)$ *given by Definition 3 is identical with the Kolmogorov-Sinaï invariant* $h(m_\varrho)$ *of the dynamical system* $(\mathscr{F} \cap \mathscr{E}(E), B, m_\varrho, T)$.

Let $B_0$ be a subfield of $B$ with finite entropy $H(B_0)$. If $M$ is any subset of $T = Z^\nu$, we define a $\sigma$ field

$$B_0(M) = \bigvee_{a \in M} \tau'_a B_0 .$$

A mean entropy $\bar{H}(B_0)$ is defined by

$$\bar{H}(B_0) = \lim_{\Lambda \to \infty} V(\Lambda)^{-1} H(B_0(\Lambda)) .$$

If $C$ is an increasingly filtered family of subfields of $B$ with finite entropy such that

$$\bigcup_{B_0 \in C} B_0(Z^\nu)$$

generates $B$ up to equivalence we know (see [4], p, 279, 6)) that the Kolmogorov-Sinaï invariant is given by

$$h = \sup_{B_0 \in C} \bar{H}(B_0) .$$

Let $\Lambda(a) = \{x_1, \ldots, x_V\}$, we define $B_a$ to be the subfield of $B$ generated by the sets $\{\varrho : \hat{A}(\varrho) > 0\}$ where $A$ is of the form $\varphi(Sf_1, \ldots, Sf_q)$ and $f_1, \ldots, f_q \in \mathscr{K}_{\Lambda(a)}$. The subsets of $\mathscr{F} \cap \mathscr{E}(E)$ obtained by specifying the numbers $n_1, \ldots, n_V$ are easily seen to be the atoms of $B_a$, hence

$$H(B_a) = S(\Lambda(a))$$

and more generally

$$H\big(B_a(\Lambda(b))\big) = S(\Lambda(a + b)) .$$

The family $C$ of all $B_a$ is clearly such that

$$\bigcup_{B_a \in C} B_a(Z^r)$$

generates $B$, and if $s(\varrho) < +\infty$, the entropies $H(B_a)$ are finite, therefore

$$h(m_\varrho) = \sup_{B_a \in C} \lim_{\Lambda(b) \to \infty} V(\Lambda(b))^{-1} H\big(B_a(\Lambda(b))\big)$$

$$= \sup_{B_a \in C} \lim_{\Lambda(b) \to \infty} \frac{V(\Lambda(a + b))}{V(\Lambda(b))} \cdot \frac{S(\Lambda(a + b))}{V(\Lambda(a + b))} = s(\varrho)$$

concluding the proof.

*Remark.* $h(m_\varrho) = +\infty$ implies $s(\varrho) = +\infty$ but we do not know if the converse holds. The resulting ambiguity would however not seem to be important in physical applications.

While one cannot expect the functional $s(\cdot)$ to be upper semi-continuous as in the continuous case, integral representations of the type given by Proposition 5 still hold. It is indeed known that the Kolmogorov-Sinaï invariant $h(m_\varrho)$ has an integral representation on $\mathscr{F} \cap \mathscr{E}(E)$ (MacMillan's theorem, [4], 10.10) and an integral representation on $\mathscr{F} \cap \mathscr{E}(E \cap \mathscr{L}^\perp)$ (barycentric decomposition, [4], 10.11).

## 5. Spin Systems

We denote by spin system a lattice system such that the occupation number $n_i$ of every lattice point $x_i$ is restricted to take the values $0$, $1, \ldots, N$ where $N < +\infty$. This terminology originates from the fact that $\frac{1}{2}(2n_i - N)$ may then be viewed as the value of a spin component (see [13]).

It is easy to construct a function $\varphi(Sf_i)$ which takes the value $0$ if $n_i \leq N$ and the value $1$ if $n_i > N$. Let $\mathscr{V}$ be the $w^*$ closed linear manifold defined by $\varphi(Sf_i) = 0$ for all $i$, the states of a spin system are then the points of $E \cap \mathscr{V}$ and we have $E \cap \mathscr{V} \subset \mathscr{F}$. The theory of spin systems is thus just a special case of the theory of lattice systems. We note however that here

$$0 \leq S(\Lambda) \leq V(\Lambda) \log(N + 1) \tag{5.1}$$

hence for all $\varrho \in E \cap \mathscr{L}^{\perp} \cap \mathscr{V}$, $s(\varrho) \in [0, \log(N+1)]$. The measure $m_\varrho$ has its support in $E \cap \mathscr{V}$ and $s(\varrho)$ is always equal to the Kolmogorov-Sinaï invariant $h(m_\varrho)$ of the dynamical system $(\mathscr{V} \cap \mathscr{E}(E),\ B,\ m_\varrho,\ T)$.

Let us now make a change of normalization in the formulas of Section 4. By writing

$$\bar{f}_A(n_1, \ldots, n_V) = (N+1)^{V(A)} f_A(n_1, \ldots, n_V) \tag{5.2}$$

$$\int \bar{d}_A\, x \cdot = (N+1)^{-V(A)} \sum_{n_1=0}^{N} \cdots \sum_{n_V=0}^{N} \cdot \tag{5.3}$$

(4.2) and (4.3) become

$$\int \bar{d}_A\, x\, \bar{f}_A(x) = 1 \tag{5.4}$$

$$\bar{f}_A(x^{(1)}) = \int \bar{d}_A\, x^{(2)}\, \bar{f}_{A'}(x^{(1)}, x^{(2)}) \tag{5.5}$$

but $\bar{d}_A\, x$ is now normalized:

$$\int \bar{d}_A\, x = 1 \tag{5.6}$$

so that we are in the same situation as for continuous systems. It follows that the entropy

$$\tilde{S}(A) = - \int \bar{d}_A\, x\, \bar{f}_A(x)\, \log \bar{f}_A(x) = S(A) - V(A) \log(N+1) \tag{5.7}$$

satisfies the inequalities of Proposition 1. Another consequence is that the affine functional $\tilde{s}(\cdot)$ is upper semi-continuous, and the same holds therefore for

$$s(\cdot) = \tilde{s}(\cdot) + \log(N+1)\,.$$

## References

1. BILLINGSLEY, P.: Ergodic theory and information. New York: John Wiley 1965.
2. CHOQUET, G., and P. A. MEYER: Ann. Inst. Fourier 13, 139 (1963).
3. DOPLICHER, S., D. KASTLER, and D. W. ROBINSON: Commun. Math. Phys. 3, 1 (1966).
4. JACOBS, K.: Lecture notes on ergodic theory. Aarhus Universitet (1962—1963).
5. KASTLER, D., and D. W. ROBINSON: Commun. Math. Phys. 3, 151 (1966).
6. LANFORD, O., and D. RUELLE: J. Math. Phys. (to appear).
7. ROBINSON, D. W., and D. RUELLE: Ann. Inst. Poincaré (to appear).
8. — Unpublished.
9. ROKHLIN, V. A.: Am. Math. Soc. Transl. 49, 171 (1966).
10. RUELLE, D.: Lecture notes of the Summer School of Theoretical Physics, Cargèse, Corsica (1965), and Commun. Math. Phys. 3, 133 (1966).
11. — J. Math. Phys. (to appear).
12. SEGAL, I. E.: Duke Math. J. 18, 221 (1951).
13. YANG, C. N., and T. D. LEE: Phys. Rev. 87, 404, 410 (1952).