Commun. math. Phys. 5, 324—329 (1967)

# A Variational Formulation of Equilibrium Statistical Mechanics and the Gibbs Phase Rule

D. Ruelle

Institut des Hautes Etudes Scientifiques
91. Bures-sur-Yvette — France

Received March 25, 1967

**Abstract.** It is shown that for an infinite lattice system, thermodynamic equilibrium is the solution of a variational problem involving a mean entropy of states introduced earlier [2]. As an application, a version of the Gibbs phase rule is proved.

## 0. Introduction

The aim of this article is to present a variational method for the determination of the equilibrium state of an infinite system in statistical mechanics. For technical reasons, we shall have to restrict ourselves to lattice systems, but it is clear that the results should extend to more general situations. As an application of the method we prove a version of the Gibbs phase rule. The physical ideas contained in this article are not different from those of an earlier paper [3], but there the programme could not be pushed through. Quite a bit of technical development has taken place since [3] which explains the ease with which the results can now be obtained. We shall in particular rely heavily on two recent papers [2] and [1] for notations and results, these are recalled in the first two sections.

## 1. Thermodynamic Limit

We consider particles on a lattice $Z^\nu$ and assume that only 0 or 1 particle can occupy a site. An interaction is a sequence $\Phi = (\Phi^{(k)})_{k \geqq 1}$ of $k$-body potentials, which are assumed to be symmetric in their arguments and invariant under translations of the lattice. Given a set $X = \{x_1, \ldots, x_n\}$ of occupied sites the potential energy is

$$U_\Phi(X) = \sum_{k \geqq 1} \sum_{1 \leqq i_1 < \cdots < i_k \leqq n} \Phi^{(k)}(x_{i_1}, \ldots, k_{i_k}) . \tag{1.1}$$

We assume that the interaction satisfies

$$\|\Phi\| = \sum_{k \geqq 1} \frac{1}{k!} \sum_{x_1, \ldots, x_{k-1} \in Z^\nu}^{\neq 0} |\Phi^{(k)}(0, x_1, \ldots, x_{k-1})| < +\infty \tag{1.2}$$

where $\sum^{\neq 0}$ extends over all sequences of distinct points of $Z^\nu$ different from 0. The interactions $\Phi$ form then a real Banach space $\mathscr{B}$ with respect to the norm (1.2), and (1.1), (1.2) give

$$|U_\Phi(X)| \leqq n\|\Phi\| . \tag{1.3}$$

For each region $\Lambda$ (finite subset of $Z^\nu$) we define a partition function

$$Z_\Lambda(\Phi) = \sum_{X \subset \Lambda} e^{-U_\Phi(X)} . \tag{1.4}$$

Let also $V(\Lambda)$ denote the number of points in $\Lambda$

**Theorem 1.** (i) *If $\Phi \in \mathscr{B}$, the following limit exists*

$$P(\Phi) = \lim_{\Lambda \to \infty} V(\Lambda)^{-1} \log Z_\Lambda(\Phi) . \tag{1.5}$$

*The functional $P(\cdot)$ is convex and continuous on $\mathscr{B}$.*

(ii) *Let $D$ be the set of all $\Phi \in \mathscr{B}$ such that the graph of $P(\cdot)$ has a unique tangent plane at $\Phi$, i. e. there is a unique element $\alpha_\Phi$ of the dual $\mathscr{B}^*$ of $\mathscr{B}$ such that for all $\psi \in \mathscr{B}$*

$$P(\Phi + \Psi) \geqq P(\Phi) - \alpha_\Phi(\Psi) . \tag{1.6}$$

*With this definition, $D$ contains a countable intersection of dense open subsets of $\mathscr{B}$ (in particular $D$ is dense by* Baire*).*

(iii) *If $\Phi \in D$, then*

$$\lim_{\Lambda \to \infty} Z_\Lambda(\Phi)^{-1} V(\Lambda)^{-1} \sum_{X \subset \Lambda} e^{-U_\Phi(X)} U_\Psi(X) = \alpha_\Phi(\Psi) . \tag{1.7}$$

*Therefore $\alpha_\Phi$ may be interpreted as the "infinite volume correlation function" corresponding to the interaction $\Phi$.*

These results are due to Gallavotti and Miracle [1].

## 2. States and their Entropy

The description of equilibrium states of systems in classical statistical mechanics has been investigated in [4] and it was shown there how such states can be identified with certain states of an adequate abelian $C^*$-algebra $\mathfrak{A}$. In the particular case of a lattice system with either 0 or 1 particle at each site, the problem is rather simple and we discuss it briefly.

Let $K$ be the product $\Pi_{x \in Z^\nu}\{0, 1\}_x$ of one copy of the set $\{0, 1\}$ (with the discrete topology) for each lattice site. With respect to the product topology, $K$ is compact. The set $E$ of states on the $C^*$-algebra $\mathscr{C}(K)$ of continuous complex functions on $K$ is naturally identical to the set of probability measures on $K$. The $v^*$-topology on $E$ is the "vague" topology of measures and makes $E$ compact. There is a bijection $\omega \to S_\omega$ of $K$ onto the subsets of $Z^\nu$ such that $\omega_x = 1 \Leftrightarrow x \in Z_\omega$. We may thus consider

$A \in \mathscr{C}(K)$ either as a function of $\omega \in K$ or of $S_\omega \subset Z^\nu$. If $x \in Z^\nu$ we define the translate $\tau_x A$ by

$$\tau_x A(S_\omega) = A(S_\omega - x) . \tag{2.1}$$

We shall be interested in the convex compact set $E \cap \mathscr{L}^\perp$ where $\mathscr{L}^\perp$ denotes the measures on $K$ which are invariant under the translations of the lattice $Z^\nu$ [i. e. $m(A) = m(\tau_x A)$]. The elements of $E \cap \mathscr{L}^\perp$ are natural candidates for the description of equilibrium states in statistical mechanics of lattice systems.

Given $\varrho \in E \cap \mathscr{L}^\perp$, for each region $\Lambda$ we define a function $f_\Lambda$ of the subsets of $\Lambda$ by

$$f_\Lambda(X) = \varrho(\{\omega \in K : S_\omega \cap \Lambda = X\}) \tag{2.2}$$

and an entropy

$$S_\varrho(\Lambda) = -\sum_{X \subset \Lambda} f_\Lambda(X) \log f_\Lambda(X) . \tag{2.3}$$

**Theorem 2.** *If $\varrho \in E \cap \mathscr{L}^\perp$, the following limit exists*

$$s(\varrho) = \lim_{\Lambda \to \infty} V(\Lambda)^{-1} S_\varrho(\Lambda) = \inf_\Lambda V(\Lambda)^{-1} S_\varrho(\Lambda) . \tag{2.4}$$

*The functional $s(\cdot)$ is affine and upper semi-continuous on $E \cap \mathscr{L}^\perp$.*

This result is contained in [2], Section 5, in a slightly different form because of a different choice of $C^*$-algebra. One may for instance derive here the upper semi-continuity of $s(\cdot)$ from the fact that $s$ is the lower bound of a family of continuous functions $V(\Lambda)^{-1} S(\Lambda)$ on $E \cap \mathscr{L}^\perp$.

## 3. The Variational Property

Given $\Phi \in \mathscr{B}$ a continuous function $A_\Phi$ on $K$ is defined by

$$A_\Phi(\omega) = \begin{cases} 0 & \text{if} \quad 0 \notin S_\omega \\ \sum_{k \geq 1} \frac{1}{k!} \sum_{x_1, \ldots, x_{k-1} \in S_\omega}^{\neq 0} \Phi^{(k)}(0, x_1, \ldots, x_{k-1}) & \text{if} \quad 0 \in S_\omega . \end{cases} \tag{3.1}$$

The linear mapping $\Phi \to A_\Phi$ is norm decreasing from $\mathscr{B}$ to $\mathscr{C}(K)$. Notice that if $S_\omega = X$ is finite

$$\sum_{x \in X} \tau_x A_\Phi(X) = U_\Phi(X) . \tag{3.2}$$

**Theorem 3** *(variational property). If $\Phi \in \mathscr{B}$, then*

$$P(\Phi) = \sup_{\varrho \in E \cap \mathscr{L}^\perp} [s(\varrho) - \varrho(A_\Phi)] . \tag{3.3}$$

We first prove that for each $\varrho \in E \cap \mathscr{L}^\perp$ we have

$$P(\Phi) \geq s(\varrho) - \varrho(A_\Phi) . \tag{3.4}$$

Let $\Lambda = \{x_1, \ldots, x_v\}$, and $\mathscr{C}_\Lambda$ be the subalgebra of $\mathscr{C}(K)$ consisting of those functions of $\omega$ which depend only on $\omega_{x_1}, \ldots, \omega_{x_v}$. If $A \in \mathscr{C}_\Lambda$, then

$$\varrho(A) = \sum_{X \subset \Lambda} f_\Lambda(X) A(X) .$$

On the other hand $U_\Lambda \mathscr{C}_\Lambda$ is dense in $\mathscr{C}(K)$ by the theorem of Stone-Weierstrass. Using this and the invariance of $\varrho$ we find that, given $\varepsilon > 0$, then for sufficiently large $\Lambda$

$$|\varrho(A_\Phi) - V(\Lambda)^{-1} \sum_{x\in\Lambda} \sum_{X\subset\Lambda} f_\Lambda(X)\, \tau_x A_\Phi(X)| < \varepsilon$$

where, by (3.2),

$$V(\Lambda)^{-1} \sum_{x\in\Lambda} \sum_{X\subset\Lambda} f_\Lambda(X)\, \tau_x A_\Phi(X) = V(\Lambda)^{-1} \sum_{X\subset\Lambda} f_\Lambda(X)\, U_\Phi(X)$$

hence

$$|\varrho(A_\Phi) - V(\Lambda)^{-1} \sum_{X\subset\Lambda} f_\Lambda(X)\, U_\Phi(X)| < \varepsilon\,. \tag{3.5}$$

On the other hand, by (2.4),

$$s(\varrho) \leqq V(\Lambda)^{-1}\Big[-\sum_{X\subset\Lambda} f_\Lambda(X) \log f_\Lambda(X)\Big]$$

therefore

$$s(\varrho) - \varrho(A_\Phi) - \varepsilon < - V(\Lambda)^{-1} \sum_{X\subset\Lambda} f_\Lambda(X)\, [U_\Phi(X) + \log f_\Lambda(X)]\,.$$

Using $\sum_{X\subset\Lambda} f_\Lambda(X) = 1$ and the concavity of the logarithm this yields

$$\begin{aligned} s(\varrho) - \varrho(A_\Phi) - \varepsilon &< V(\Lambda)^{-1} \sum_{X\subset\Lambda} f_\Lambda(X) \log \frac{e^{-U_\Phi(X)}}{f_\Lambda(X)} \leqq \\ &\leqq V(\Lambda)^{-1} \log \sum_{X\subset\Lambda} e^{-U_\Phi(X)}\,. \end{aligned}$$

By (1.5), this yields (3.4) when $\Lambda \to \infty$.

We show now that $\varrho \in E \cap \mathscr{L}^\perp$ can be found such that

$$P(\Phi) < s(\varrho) - \varrho(A_\Phi) + 2\varepsilon\,. \tag{3.6}$$

Given an integer $n > 0$, let

$$\Lambda_n = \{x \in Z^\nu : 0 \leqq x^i < n \quad \text{for} \quad i = 1, \ldots, \nu\}\,.$$

For $n$ large enough, (1.5) yields

$$|P(\Phi) - V(\Lambda_n)^{-1} \log Z_{\Lambda_n}(\Phi)| < \varepsilon\,. \tag{3.7}$$

We introduce a function $\bar{f}$ of finite subsets of $Z^\nu$ by

$$\bar{f}(X) = Z_{\Lambda_n}(\Phi)^{-1}\, e^{-U_\Phi(X)}\,.$$

The translates

$$\Lambda_n + n\,k$$

of $\Lambda_n$, where $k \in Z^\nu$ form a partition of $Z^\nu$. Let $\Lambda$ be the union of a finite number of such translates

$$\Lambda = \bigcup_{i=1}^{N} [\Lambda_n + n\,k_i]\,.$$

If $A \in \mathscr{C}_\Lambda$, we define

$$\tilde{\varrho}(A) = \sum_{X_1\subset\Lambda_n + n k_1} \cdots \sum_{X_N\subset\Lambda_n + n k_n} \bar{f}(X_1) \ldots \bar{f}(X_N)\, A(X_1 \cup \cdots \cup X_N)\,.$$

It is easy to see that this definition does not depend on the special choice of $\Lambda$ and, the union of the $\mathscr{C}_\Lambda$ being dense in $\mathscr{C}(K)$, $\tilde{\varrho}$ extends uniquely to a state on $\mathscr{C}(K)$, which is periodic with periodicity cell $\Lambda_n$. A state $\varrho \in E \cap \mathscr{L}^\perp$ is now obtained by averaging $\tilde{\varrho}$ over translations:

$$\varrho(A) = V(\Lambda_n)^{-1} \sum_{x \in \Lambda_n} \varrho(\tau_x A)$$

then, by easy estimates

$$\begin{aligned} s(\varrho) &= \lim_{\Lambda \to \infty} V(\Lambda)^{-1} \big[-\sum_{X \subset \Lambda} f_\Lambda(X) \log f_\Lambda(X)\big] \\ &= V(\Lambda_n)^{-1} \big[-\sum_{X \subset \Lambda_n} \tilde{f}(X) \log \tilde{f}(X)\big]. \end{aligned} \tag{3.8}$$

Using for instance (3.5) one checks also that for large $n$

$$|\varrho(A_\Phi) - V(\Lambda_n)^{-1} \sum_{X \subset \Lambda_n} \tilde{f}(X)\, U_\Phi(X)| < \varepsilon\,. \tag{3.9}$$

From (3.8), (3.9) and (3.7) we get

$$\begin{aligned} s(\varrho) - \varrho(A_\Phi) + \varepsilon &> V(\Lambda_n)^{-1} \big[-\sum_{X \subset \Lambda_n} \tilde{f}(X)\, [U_\Phi(X) + \log \tilde{f}(X)]\big] \\ &= V(\Lambda_n)^{-1} \log Z_{\Lambda_n}(\Phi) > P(\Phi) - \varepsilon \end{aligned}$$

which proves (3.6) and therefore the Theorem.

## 4. The Gibbs Phase Rule

Given a one-component thermodynamic system, we take the Gibbs phase rule to mean that "almost all" points of the $(\mu, \beta)$ diagramme correspond to a single phase, $\mu$ being the chemical potential and $\beta$ the inverse temperature. It is however conceivable that for special choices of the interaction this statement becomes incorrect. One is thus led to formulate the following "Gibbs phase rule": for "almost all" interactions and $(\mu, \beta)$ a system at equilibrium consists of only one phase. We deal with classical systems and $\mu$ may be considered as a "one-body potential" while $\beta$ is a multiplicative factor for all potentials. We may thus omit $\mu$, $\beta$ in the formulation of the Gibbs phase rule (absorbing $\mu$ in the interaction and putting $\beta = 1$). We say that a system consists of only one phase when it is described by a state $\varrho$ which is an extremal point of $E \cap \mathscr{L}^\perp$ (see [3], [4]), i. e. an ergodic measure on $K$. The Gibbs phase rule means thus that ergodicity is generic for the states of infinite systems in classical statistical mechanics. To obtain a theorem it remains only to precise the set of interactions considered and the notion of "almost all" interactions in this set.

**Theorem 4** *(Gibbs phase rule).* (i) *If $\Phi$ belongs to the set $D$ of* Theorem 1, *the function $\varrho \to s(p) - \varrho(A_\Phi)$ reaches its maximum $P(\Phi)$ at exactly one point $\varrho_\Phi \in E \cap \mathscr{L}^\perp$.*

(ii) *If* $\Phi \in D$ *and* $\alpha_\Phi$ *is the functional defined in* Theorem 1 *then, for all* $\Psi$,

$$\varrho_\Phi(A_\Psi) = \alpha_\Phi(\Psi) \tag{4.1}$$

*and* $\varrho_\Phi$ *may be interpreted as the "equilibrium state" corresponding to the interaction* $\Phi$.

(iii) *If* $\Phi \in D$, *the equilibrium state* $\varrho_\Phi$ *is an extremal point* $E \cap \mathscr{L}^\perp$ *(= ergodic measure on* $K$ *= pure thermodynamic phase).*

For any $\Phi \in \mathscr{B}$, the functional $\varrho \to s(\varrho) - \varrho(A_\Phi)$ is affine upper semi-continuous on $E \cap \mathscr{L}^\perp$ and reaches thus its maximum $P(\Phi)$ on a non-empty set $\Delta_\Phi$ which is convex and compact and contains therefore at least one extremal point of $E \cap \mathscr{L}^\perp$. If $\varrho \in \Delta_\Phi$, we have for all $\Psi \in \mathscr{B}$

$$P(\Phi+\Psi) \geqq s(\varrho) - \varrho(A_{\Phi+\Psi}) = s(\varrho) - \varrho(A_\Phi) - \varrho(A_\Psi) = P(\Phi) - \varrho(A_\Psi). \tag{4.2}$$

Since $\Psi \to A_\Psi$ is linear and continuous, (4.2) shows that $\Psi \to P(\Phi) - \varrho(A_\Psi)$ is a tangent plane to the graph of $P(\cdot)$ at $\Phi$. Let $\mathscr{L}$ be the Banach subspace of $\mathscr{C}(K)$ generated by differences $A - \tau_x A$ with $A \in \mathscr{C}(K)$, $x \in Z^\nu$, and $Q : \mathscr{C}(K) \to \mathscr{C}(K)/\mathscr{L}$ be the quotient mapping. We notice that $\mathscr{L}^\perp$ (invariant measures on $K$) is isomorphic as Banach space to the dual of $\mathscr{C}(K)/\mathscr{L}$, and that $\{Q A_\Psi : \Psi \in \mathscr{B}\}$ is dense in $\mathscr{C}(K)/\mathscr{L}$ (use the fact that $U_\Lambda \mathscr{C}_\Lambda$ is dense in $\mathscr{C}(K)$). Therefore if an element of $\mathscr{L}^\perp$ vanishes on all $A_\Psi$ it vanishes identically, in particular two distinct elements of $\Delta_\Phi$ yield different tangent planes to the graph of $P(\cdot)$. By theorem 1, if $\Phi \in D$, $\Delta_\Phi$ is thus reduced to one point $\varrho_\Phi$ which is extremal in $E \cap \mathscr{L}^\perp$ and $\varrho_\Phi(A_\Psi) = \alpha_\Phi(\Psi)$ for all $\Psi \in \mathscr{B}$.

*Remark.* The variational formulation of equilibrium given in this paper is grand canonical. In [3] a microcanonical approach was used: "equilibrium is realized by the state which, for a given density and energy density (with respect to a given interaction), has maximum entropy". As usual the microcanonical point of view is physically more intuitive, the grand canonical point of view is technically easier to handle.

We want also to point out that the reader can reintroduce the temperature and chemical potential explicitly (see [1]), he will find in particular that for a dense set of interactions in $\mathscr{B}$, the set of $(\mu, \beta)$ points for which there is more than one phase is of Lebesgue-measure zero.

## Bibliography


1. Gallavotti, G., and S. Miracle-Sole: Statistical mechanics of lattice systems. Commun. Math. Phys. **5**, 317 (1967).
2. Robinson, D. W., and D. Ruelle: Mean entropy of states in classical statistical mechanics. Commun. Math. Phys. **5**, 288 (1967).
3. Ruelle, D.: J. Math. Phys. **6**, 201—220 (1965).
4. — The states of classical statistical mechanics. Preprint.