Journal of Statistical Physics, Vol. 20, No. 5, 1979

The Equivalence of Ensembles and the Gibbs Phase Rule for Classical Lattice Systems

Anders Martin-Löf[1]

Received October 25, 1977; revised November 15, 1978

We study the relation between the microcanonical, canonical, and grand canonical ensembles in the thermodynamic limit when the system becomes infinite. They are equivalent if there is only one phase in the system. In general it is shown that there is a unique limit of the microcanonical state being a mixture of pure phases if the microcanonical restrictions determine the volume fractions of the phases uniquely, and then the Gibbs phase rule is valid. In this context we show how to define the set of order parameters associated with the state of the system in a natural way.

KEY WORDS: Equivalence of different ensembles; order parameters; equilibrium states; pure phases; variational principles; convexity.

1. INTRODUCTION

In this paper we study the relation between a restricted, microcanonical, and the corresponding unrestricted, grand canonical, probability law (ensemble, state) for a system in a finite container described by classical statistical mechanics. We want to investigate how these ensembles are related in the thermodynamic limit when the volume tends to infinity and the values of the variables which are restricted are proportional to the volume. The result indicated heuristically in many textbooks that for a large system described by a microcanonical law the state of a small subsystem is given by the corresponding grand canonical law is proved in the following sense: Consider the state of a small subvolume whose position is chosen at random with uniform distribution in the large container. Then if there is no phase transition for the given values of the restricted variables this state has a unique limit equal to the unique invariant equilibrium state defined by the parameters of the grand canonical state, and it is also the limit of the latter

[1] University of Stockholm, Department of Mathematics, Stockholm, Sweden.

0022-4715/79/0500-0557$03.00/0

state. If there is a phase transition, so there is more than one invariant equilibrium state, each being a mixture of extremal ones, i.e., states describing the pure phases, then we have the following situation: There is a unique limit of the state of the subvolume being a mixture of the states of the pure phases if the microcanonical restrictions determine the mixture completely, and there is only a finite number of pure phases. In this case this number is at most one plus the number of restrictions (Gibbs phase rule.)

This indicates that the weights in the mixture are to be interpreted as the fractions of the total volume filled by the various pure phases, because if these occupy macroscopic regions the probability that the subvolume falls inside a region filled by a pure phase is proportional to its volume fraction, and the probability that it overlaps with several phases is negligible when the volume is large.

We also show that in this context one obtains a natural definition of the order parameters associated with the phase transition. One can also study the limit of a mixed, canonical, ensemble, where some of the restrictions are relaxed and taken care of by the grand canonical parameters. It then follows that any restrictions not defined by the order parameters can be relaxed without changing the unique limit of the microcanonical state. For simplicity we consider only classical lattice systems for which the theory of invariant equilibrium states is well developed. Our proofs are based on the variational characterization of these and on general arguments based on the convexity properties of the entropy and pressure, so they are certainly valid for more general systems.

2. BASIC DEFINITIONS AND FACTS

We consider a classical lattice system on $\mathbb{Z}^d = L$ with a finite set S of possible states at each lattice site, i.e., the configuration space is $X = S^L$. Configurations are denoted by x, y, etc. and x_A denotes the restriction of $x \in X$ to the subset $A \subset L$. We also consider configurations defined only on subsets of L, and for such a configuration x, $\mathbf{x}$ denotes its domain of definition. X_A is the set of x with $\mathbf{x} = A$, $A \subset L$, and X_f is the set of all x with finite domain, $|\mathbf{x}| < \infty$. For any $p \in L$, $T_p x$ denotes the configuration obtained by translating x by p.

A potential is a translation-invariant function $U(x)$ defined for $x \in X_f$. For Λ finite $\subset L$ and $x \in X_\Lambda$ the total "U-energy" of x is

$$U_\Lambda(x) = \sum_{A \subseteq \Lambda} U(x_A)$$

We consider as is usual the Banach space B of potentials with finite norm defined by

$$\|U\| = \sum_{A \ni 0} |A|^{-1} \sup_{x \in X_A} |U(x)|$$

in which the set B_f of potentials with finite range is dense. U is of finite range if and only if there is a finite set $R \subset L$ so that $U(x) = 0$ unless $\mathbf{x} \subset T_p R$ for some $p \in L$.

In terms of $\|U\|$ we have the following bound for $U_\Lambda(x)$:

$$|U_\Lambda(x)| \leqslant \sum_{A \subseteq \Lambda} |U(x_A)| = \sum_{A \subseteq \Lambda} \sum_{p \in A} |A|^{-1} |U(x_A)|$$
$$\leqslant \sum_{p \in \Lambda} \sum_{p \in A} |A|^{-1} |U(x_A)| \leqslant |\Lambda| \cdot \|U\|$$

so $|\Lambda|^{-1} U_\Lambda(x) \leqslant \|U\|$ for all Λ, and $x \in X_\Lambda$.

A restricted probability law for a finite container Λ is defined by giving a finite number of potentials $U_1, \ldots, U_m \in B$ and giving equal probability to all $x \in X$ having nearly fixed values of $U_{1\Lambda}(x), \ldots, U_{m\Lambda}(x)$:

Let I be a small, open neighborhood of a point $u \in R^m$, and let $Z_\Lambda(I) =$ number of $x \in X_\Lambda$ such that $|\Lambda|^{-1} U_\Lambda(x) \in I$. Then the restricted law is given by

$$p_\Lambda(x; I) = \begin{cases} Z_\Lambda^{-1}(I) & \text{if } |\Lambda|^{-1} U_\Lambda(x) \in I \\ 0 & \text{otherwise} \end{cases}$$

$$U_\Lambda(x) = (U_{1\Lambda}(x), \ldots, U_{m\Lambda}(x)) \in R^m, \qquad a \cdot U = \sum_1^m a_i U_i, \quad \text{etc.}$$

We consider a "macroscopically thin energy shell" defined by $|\Lambda|^{-1} U_\Lambda(x) \in I$, and we will consider the limit $\Lambda \to \infty$ and then $I \to u \in R^m$. This is physically natural and mathematically simpler than having $\Lambda \to \infty$, $|\Lambda|^{-1} U_\Lambda(x) \to u$ simultaneously.

For definiteness we always take Λ to be a parallelepiped, and $\Lambda \to \infty$ means that all the sides of Λ go to infinity.

The unrestricted probability law corresponding to the one defined above is defined by introducing m intensive parameters $(a_1, \ldots, a_m) = a \in R^m$ and relaxing the constraints:

$$p_\Lambda(x; a) = Z_\Lambda^{-1}(a) \exp[-a \cdot U_\Lambda(x)] \qquad \text{for} \quad x \in X_\Lambda$$

with

$$Z_\Lambda(a) = \sum_{x \in X_\Lambda} \exp[-a \cdot U_\Lambda(x)] = \int \exp[-|\Lambda|(a \cdot u)]\, Z_\Lambda(du)$$

In general, for any $V \in B$, not necessarily of the form $a \cdot U$, we also define

$$p_\Lambda(x; V) = Z_\Lambda^{-1}(V) \exp[-V_\Lambda(x)] \qquad \text{for} \quad x \in X_\Lambda$$

with

$$Z_\Lambda(V) = \sum_{x \in X_\Lambda} \exp[-V_\Lambda(x)]$$

[We make the abuse of notation $Z_\Lambda(a) = Z_\Lambda(a \cdot U)$, etc., when there is no risk of confusion.]

As a functional on B, $\log Z_\Lambda(V)$ is a generating functional for the state of subsystems located at random with a uniform distribution over Λ as follows: Put $g_\Lambda(V) = |\Lambda|^{-1} \log Z_\Lambda(V)$; then for $U \in B$

$$-g_\Lambda'(V, U) \equiv -\frac{d}{d\lambda} g_\Lambda(V + \lambda U) \Big|_{\lambda=0} = \langle |\Lambda|^{-1} U_\Lambda(x) \rangle_\Lambda$$
$$= |\Lambda|^{-1} \sum_{A \subseteq \Lambda} \langle U(x_A) \rangle_\Lambda = |\Lambda|^{-1} \sum_{y \subseteq \Lambda} U(y) p_\Lambda(x_y = y; V)$$

Put

$$p_\Lambda(y; V) = \begin{cases} p_\Lambda(x_y = y; V) & \text{if } y \subseteq \Lambda \\ 0 & \text{otherwise} \end{cases}$$

(These are the "cylinder probabilities" of the state on X_Λ. Any state on X is uniquely specified by its cylinder probabilities for $y \in X_f$.)

Using the same change of summation as above, we then get

$$-g_\Lambda'(V, U) = |\Lambda|^{-1} \sum_{p \in \Lambda} \sum_{p \in A} |A|^{-1} \sum_{y=A} U(y) p_\Lambda(y; V)$$
$$= |\Lambda|^{-1} \sum_{p \in \Lambda} \sum_{0 \in B} |B|^{-1} \sum_{y = T_p B} U(y) p_\Lambda(y; V)$$
$$= \sum_{0 \in B} |B|^{-1} \sum_{x = B} |\Lambda|^{-1} \sum_{p \in \Lambda} U(x) p_\Lambda(T_p x; V)$$
$$= \sum_{x \ni 0} |x|^{-1} U(x) \bar{p}_\Lambda(x; V) = (\bar{p}_\Lambda(\cdot\,; V), U)$$

if we put

$$\bar{p}_\Lambda(x; V) \equiv |\Lambda|^{-1} \sum_{p \in \Lambda} p_\Lambda(T_p x; V)$$
$$(\bar{p}_\Lambda, U) \equiv \sum_{x \ni 0} |x|^{-1} \bar{p}_\Lambda(x) U(x)$$

For any finite $A \ni 0$ the $\bar{p}_\Lambda(x; V)$ with $x \in X_A$ describe the probability law for the state in a subvolume $T_p A$ chosen at random in Λ, so they are the basic probabilities we want to study. When $V = a \cdot U$, a and u are said to correspond if the unrestricted law is centered so that $\langle |\Lambda|^{-1} U_\Lambda(x) \rangle = u$, i.e., if $-g_\Lambda'(a) = u$.

Now, $g_\Lambda(V)$ is convex in V, and hence $g_\Lambda'(V, U)$ defines a subgradient to g_Λ at V:

$$g_\Lambda(V + U) - g_\Lambda(V) \geqslant g_\Lambda'(V, U) = -(\bar{p}_\Lambda(\cdot\,; V), U) \qquad \text{for all} \quad U \in B$$

We recall the following basic facts about the thermodynamic limit of g_Λ (Ref. 9): $\lim_{\Lambda \to \infty} g_\Lambda(V) = g(V)$ exists and is a bounded, convex function on B; $g_\Lambda(V) \leqslant \|V\| + \log|S|$; and $|g_\Lambda(V) - g_\Lambda(U)| \leqslant \|V - U\|$. Any limit $\bar{p}(x; V)$ of the $\bar{p}_\Lambda(x; V)$, $x \in X_f$ for some subsequence $\Lambda \to \infty$, defines a subgradient to g at V, and hence if g is differentiable at V, so there is only one such subgradient, all such limits are the same, i.e., $\lim_{\Lambda \to \infty} \bar{p}_\Lambda(x; V) =$

$\bar{p}(x; V)$ exists for all $x \in X_f$. The $\bar{p}(x; V)$ define cylinder probabilities for an invariant state on X.

We also recall the following facts about invariant equilibrium states:[3,9] The unrestricted ensemble in a finite set Λ can be characterized by the Gibbs variational principle: For any probability distribution p_Λ on X_Λ let

$$H(p_\Lambda) = -\sum_{x \in X_\Lambda} p_\Lambda(x) \log p_\Lambda(x)$$

Then $p_\Lambda(x; V)$ is the distribution that maximizes

$$H(p_\Lambda) - \langle V_\Lambda(x) \rangle_\Lambda = H(p_\Lambda) - \sum_{x \in X_\Lambda} V_\Lambda(x) p_\Lambda(x)$$

and the value of the maximum is $\log Z_\Lambda(V)$. An invariant equilibrium state on X defined by invariant cylinder probabilities $p(x)$, $x \in X_f$, is characterized by an analogous maximization:

$$h(p) = \lim_{\Lambda \to \infty} -|\Lambda|^{-1} \sum_{x \in X_\Lambda} p(x) \log p(x)$$

and

$$\lim_{\Lambda \to \infty} |\Lambda|^{-1} \sum_{x \in X_\Lambda} p(x) V_\Lambda(x) = (p, V)$$

are always defined for any invariant state, and $\max_p(h(p) - (p, V))$ is always attained for some p and equal to $g(V)$. The maximal p's are called invariant equilibrium states and form a compact convex set of invariant states which is a Choquet simplex, so that any such state can be uniquely represented as a mixture of extremal such states:

$$p = \int_{E_V} p_e \, q(de)$$

These extremal states p_e are ergodic and describe "pure phases" for the system. Any equilibrium state p for V defines a subgradient to g at V:

$$\begin{aligned} g(U) - g(V) &= \max_p(h(p) - (p, U)) - h(p) + (p, V) \\ &\geqslant h(p) - (p, U) - h(p) + (p, V) = -(p, U - V) \end{aligned}$$

and for our purposes it will be important to know that any subgradient to g at V defines an invariant equilibrium state, so these are precisely the subgradients to g at V.[3]

Example. Our basic example is the Ising model, the only one for which one knows precisely which are the equilibrium states for all parameter values. It has $S = \{-1, +1\}$, $m = 2$, and $U_{1\Lambda}(x)$ is the total interaction energy, $U_{2\Lambda}(x)$ is the total magnetization. We have $U_1(x) \neq 0$ only if $\mathbf{x}$ is a pair of nearest neighbors $\{p, q\}$ on L and then $U_1(x) = x_p \cdot x_q$. Now, $U_2(x) \neq 0$ only if $\mathbf{x}$ is a single point $p \in L$, and then $U_2(x) = x_p$. For the

unrestricted ensemble $(a_1, a_2) = (\beta, -\beta h)$, where β is the inverse temperature $(kT)^{-1}$ and h the external field. In Refs. 4 and 6 it is shown for $d = 2$ that for $h \neq 0$ there is precisely one invariant equilibrium state and also for $h = 0$, $\beta \leqslant \beta_c$, β_c being the critical temperature. For $\beta > \beta_c$ there are precisely two extremal invariant equilibrium states $p_\pm$ obtained from the unique state for $h \neq 0$ by passing to the limit $h \to \pm 0$, respectively. The $p_\pm$ are related by the symmetry $x \to -x$ and have average magnetization $\pm m^*(\beta) \neq 0$, respectively.

3. THE THERMODYNAMIC LIMIT FOR THE RESTRICTED ENSEMBLE

The restricted probability law was expressed in terms of $Z_\Lambda(I)$, $I \subset R^m$, defined above. Let us more generally define $Z_\Lambda(I, a)$ as

$$Z_\Lambda(I, a) = \sum_{\substack{|\Lambda|^{-1} U(x) \in I \\ x \in X_\Lambda}} \exp[-a \cdot U_\Lambda(x)]$$

Then we have $Z_\Lambda(I, 0) = Z_\Lambda(I)$ and $Z_\Lambda(R^m, a) = Z_\Lambda(a)$. The thermodynamic limit of $Z_\Lambda(I, a)$ can be described as follows: $\lim_{\Lambda \to \infty} |\Lambda|^{-1} \log Z_\Lambda(I, a) = s(I, a)$ exists if I is open convex $\subset R^m$ (or a finite union of such sets). $s(I, a)$ is bounded above if I is bounded, but can be equal to $-\infty$. It has the following form: $s(I, a) = \sup_{u \in I}(s(u) - a \cdot u)$ with $s(u) = \inf_{I \ni u} s(I, 0)$. The entropy function $s(u)$ is concave, upper semicontinuous (u.s.c.), and bounded above by $\log|S|$. It is the unique concave u.s.c. function representing $s(I, a)$ in the above fashion. (The proofs of these facts are given in detail in Refs. 3 and 5. Actually a stronger condition on the potential than $U_i \in B$ is used there: $U_i \in B_f$ or U_i being "tempered," but probably this is not necessary.)

In particular for $I = R^m$ we get

$$s(R^m, a) = g(a) = \sup_u(s(u) - a \cdot u)$$

so s and g are "conjugate functions" or Legendre transforms of each other. We recall the following facts about this correspondence:[5,8]

$$s(u) = \inf_a(g(a) + a \cdot u)$$

$g(a)$ is the minimal g such that the affine function $g + a \cdot u \geqslant s(u)$ for all u (Fig. 1) and $s(u)$ is the maximal s such that $s - a \cdot u \geqslant g(a)$ for all a (Fig. 2). Now, a and u correspond in $\sup_u$, i.e., $\sup_u$ is attained at u iff they correspond in $\inf_a$, and this happens iff they define supporting planes to $s(u)$ and $g(a)$, respectively, as in the figures. There is a unique u corresponding to a iff s has no linear segment with slope a in its graph. Conversely there is a unique a corresponding to u iff s is differentiable at u, and then $a = s'(u)$, or iff g has no linear segment in its graph with slope $-u$.

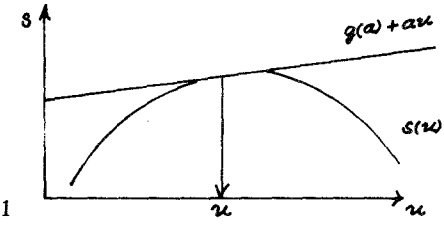

Fig. 1

In Ref. 2 it is shown that $g(V)$ is strictly convex in V so $g(a) = g(a \cdot U)$ is also strictly convex, and there is hence a unique a corresponding to u if $s(u) > -\infty$ if the U_i are linearly independent. Since we have seen that $Z_\Lambda(I) = 0$ if I is disjoint from the compact region where $|u_i| \leqslant \|U_i\|$, $i = 1,..., m$, then $s(u) = -\infty$ outside of this region. This means that $\sup_u$ is always attained for some u, because an u.s.c. function always takes on a maximal value on a compact set.

Let us now see how the entropy function can be used to study the restricted ensemble in the limit $\Lambda \to \infty$, $I \to u$. Suppose that we want to study the distribution of some other variable $U_{0\Lambda}(x)$ for $U_0 \in B_f$, e.g. Then we include U_0 among $U_1,..., U_m$ and define $Z_\Lambda(I_0 \times I)$, etc., as before by the restriction $|\Lambda|^{-1}(U_{0\Lambda}(x), U_\Lambda(x)) \in I_0 \times I$, $I_0 \subset R^1$, $I \subset R^m$, and get

$$s(I_0 \times I) = \sup_{u_0 \in I_0; u \in I} s(u_0, u)$$

[We use the same symbol for the extended function $s(u_0, u)$, etc., without risk of confusion.] Then we have

$$p_\Lambda(|\Lambda|^{-1} U_{0\Lambda}(x) \in I_0; I) \equiv p_\Lambda(I_0; I) = \frac{Z_\Lambda(I_0 \times I)}{Z_\Lambda(R^1 \times I)} = \frac{Z_\Lambda(I_0 \times I)}{Z_\Lambda(I)}$$

so that as $\Lambda \to \infty$

$$|\Lambda|^{-1} \log p_\Lambda(I_0; I) \to s(I_0 \times I) - s(R^1 \times I) = s(I_0 \times I) - s(I)$$

when $s(I) > -\infty$. It then follows that as $I \to u$

$$\lim_{I \to u} \lim_{\Lambda \to \infty} |\Lambda|^{-1} \log p_\Lambda(I_0; I) = \sup_{u_0 \in I_0} s(u_0, u) - \sup_{u_0} s(u_0, u)$$
$$\equiv s(I_0, u) - s(R^1, u)$$

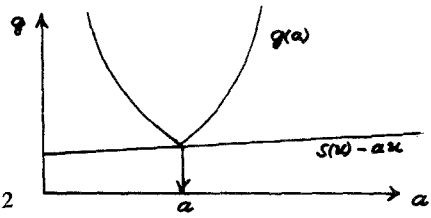

Fig. 2

when

$$s(u) = \sup_{u_0} s(u_0, u) > -\infty$$

[Consider, e.g., the first term as a function of I:

$$s(I_0 \times I) = \sup_{u \in I}(\sup_{u_0 \in I_0} s(u_0, u)) = \sup_{u \in I} s(I_0, u)$$

Hence $s(I_0, \cdot)$ represents $s(I_0 \times \cdot)$, so if it is u.s.c. it will be given by $\inf_{I \in u} s(I_0, I)$. This is easy to check using the fact that $s(u_0, u) = -\infty$ when $|u_0| > \|U_0\|$.(5)]

We now see directly that all the probability mass of the distribution $p_\Lambda(I_0; I)$ will be concentrated at those values of u_0 for which $\sup_{u_0} s(u_0, u)$ is attained: Let $M = [u_0', u_0'']$ be this maximal interval. Then if $\bar{I}_0$ is disjoint from M, $s(I_0, u) - s(R^1, u) < 0$, so

$$\lim_{\Lambda \to \infty} |\Lambda|^{-1} \log p_\Lambda(I_0; I) < 0 \quad \text{if} \quad I \approx u \quad \text{and} \quad \lim_{\Lambda \to \infty} p_\Lambda(I_0; I) = 0$$

Since $|\Lambda|^{-1} U_{0\Lambda}(x)$ is bounded, this also means that any limiting value of $\langle |\Lambda|^{-1} U_{0\Lambda}(x) \rangle_\Lambda$ must be contained in M. In particular if M consists of only one point u_0, this average converges to u_0 as $\Lambda \to \infty$, $I \to u$. It also follows that any limit $p(x)$, $x \in X_f$, of the cylinder probabilities $\bar{p}_\Lambda(x; I)$ defines an equilibrium state (Such limits always exist by the compactness of the set of states on X.)

Lemma 1. If a corresponds to u in the duality, i.e., if $g(a) = s(u) - a \cdot u$ and $V = a \cdot U$, then any limit $p(x)$, $x \in X_f$, of the $\bar{p}_\Lambda(x; I)$ defines a subgradient to g at V and hence an equilibrium state for V.

Proof. What has to be shown is that, for any a_0 and U_0, $g(a \cdot U + a_0 U_0) - g(a \cdot U) \leqslant -a_0(p, U_0)$, i.e., $g(a_0, a) - g(0, a) \geqslant -a_0 u_0$, $u_0 = (p, U_0)$.

This follows because we have $g(a_0, a) \geqslant s(u_0, u) - a_0 u_0 - a \cdot u$, $g(0, a) = g(a) = s(u) - a \cdot u$, and $s(u) = s(u_0, u)$ since $\sup_{u_0} s(u_0, u)$ is attained at such a u_0.

Hence if there is no phase transition for $V = a \cdot U$ so g is differentiable at V and there is a unique equilibrium state, this is equal to the limit of both the restricted and the unrestricted states. I.e., we have equivalence of ensembles when there is no phase transition.

We now consider the case when there are several equilibrium states for $V = a \cdot U$. As we have just seen, any limit $p(x)$ of the $\bar{p}_\Lambda(x; I)$ defines an equilibrium state and is hence of the form $p = \int_{E_V} p_e \, q(de)$, and we have

$$u_i = \lim \langle |\Lambda|^{-1} U_i(x) \rangle_\Lambda = (p, U_i) = \int_{E_V} (p_e, U_i) \, q(de) \quad \text{for} \quad i = 1, \ldots, m$$

Hence if these equations determine $q(de)$ uniquely, $p(x)$ is uniquely determined and equal to the limit of $\bar{p}_\Lambda(x; I)$ for all $x \in X_f$. This is our main result:

Theorem 1. Given $V = a \cdot U$. Then the following statements are equivalent:

A. For all u corresponding to a in the duality and for all $U_0 \in B_f$, $s(u_0, u)$ has a unique u_0 giving $\sup_{u_0} s(u_0, u)$.

B. There are only finitely many pure phases for V, $\{p_e, e = 1,..., f\}$, and the equations

$$u_i = \sum_1^f (p_e, U_i) q_e, \quad i = 1,..., m; \qquad 1 = \sum_1^f q_e$$

determine the mixture $\{q_e\}$ uniquely for all u corresponding to a.

If B holds, then $f \leqslant m + 1$ (Gibbs phase rule) and

$$\lim_{I \to u} \lim_{\Lambda \to \infty} \bar{p}_\Lambda(x; I) = \sum_1^f p_e(x) q_e, \qquad x \in X_f$$

Proof. A $\Rightarrow$ B: If the equations

$$u_i = \int_{E_V} (p_e, U_i)\, q(de), \quad i = 1,..., m; \qquad 1 = \int_{E_V} q(de)$$

had two different solutions q', q'' for some u corresponding to a, then these would define different equilibrium states or subgradients p', p'' to g at V. I.e., for some $U_0 \in B_f$ we would have $u_0' = (p', U_0) < (p'', U_0) = u_0''$ and

$$g(a_0 U_0 + b \cdot U + V) - g(V) \geqslant -a_0 (p, U_0) - b \cdot (p, U)$$

both for p', ρ'', i.e.,

$$g(a_0, b + a) - g(0, a) \geqslant -a_0 u_0 - b \cdot u$$

both for u_0', u_0''. I.e., both (u_0', u) and (u_0'', u) would correspond to $(0, a)$ and $s(u_0, u)$ would be maximal both for u_0', u_0'' against the assumption. We hence see that $\{u; u_i = \int_{E_V} (p_e, U_i)\, q(de)\}$ is a simplex in R^m with extreme points (p_e, U), $e \in E_V$. Hence E_V has to be finite with at most $f \leqslant m + 1$ elements.

B $\Rightarrow$ A: Any u_0 such that $s(u_0, u)$ is maximal as we have seen defines a subgradient $-(u_0, u)$ to g at $(0, a)$, i.e.,

$$g(a_0 U_0 + b \cdot U + V) - g(V) \geqslant -a_0 u_0 - b \cdot u$$

i.e., we have a subgradient in the subspace of B spanned by $(U_0,..., U_m)$. By the Hahn–Banach theorem this can be extended to a subgradient to g at V in all of B, i.e., there is an equilibrium state $p = \sum_1^f p_e q_e$ such that $(p, U_i) = u_i$, $i = 0, 1,..., m$. The equations for $i = 1,..., m$ determine the q_e uniquely, and then u_0 is uniquely given by $u_0 = \sum_1^f (p_e, U_0) q_e$.

4. THE DEFINITION OF THE ORDER PARAMETERS

Let us consider the structure of the equations determining the mixture:

$$u_i = \sum_1^f (p_e, U_i) q_e, \quad i = 1,..., m; \qquad 1 = \sum_1^f q_e$$

We assume without loss of generality that the U_i are linearly independent. If, for some i, $(p_e, U_i) = u_i$ for all e, then the corresponding equation in the above system is redundant for determining the q_e, so only the remaining equations serve this purpose. This happens iff g is differentiable in the direction U_i at $V = a \cdot U$ according to the following lemma:

Lemma 2. g is differentiable in a direction W at V iff (p, W) has a constant value w for all subgradients to g at V, and then $w = -g'(V, W)$. Hence these directions form a linear subspace, D, of B.

Proof. If $g'(V, W)$ is defined and $-p$ is a subgradient, we have

$$g(V + bW) - g(V) \geqslant -b(p, W) \qquad \text{for all } b$$

and hence $(p, W) = -g'(V, W)$.

Conversely, if $-w$ is any subgradient in the direction W, i.e., if $g(V + bW) - g(V) \geqslant -bw$ for all b, then the Hahn–Banach result tells us that there is a subgradient $-p$ with $w = (p, W)$. Hence w can take only one value if (p, W) is independent of p. Hence $-w_\pm = g_\pm'(V, W)$, the right and left derivatives in the direction W are equal because they both define subgradients in the direction W, and $g'(V, W)$ is defined and equal to their common value.

In the space spanned by $U_1,..., U_m$ we can then suppose that, e.g., $U_{n+1},..., U_m$ span its intersection with D, so that $(p_e, U_i) = u_i$ for all e and $i = n+ 1,..., m$, and only the remaining equations for $i = 1,..., n$ serve to determine the q_e. These equations are linearly independent, since if we had $\sum_1^n c_i(p_e, U_i) + c = 0$ for all e we would have $\sum_1^n c_i U_i \in D$, and $U_{n+1},..., U_m$ would not be a basis for the intersection with D. Hence the equations have full rank and $f = n + 1$.

The variables $U_{1\Lambda}(x),\ldots, U_{n\Lambda}(x)$ should be called the "order parameters" at $V = a \cdot U$, i.e., those variables that in addition to a have to be fixed in order that the state should be completely determined. (They are of course only determined up to linear equivalence.)

5. THE THERMODYNAMIC LIMIT FOR THE MIXED ENSEMBLE

It is also of interest to consider an intermediate ensemble where some variables $U_{1\Lambda}(x),\ldots, U_{m\Lambda}(x)$ take nearly fixed values and the probability density for $x \in X_\Lambda$ is proportional to $\exp[-b W_\Lambda(x)]$ for some other variable $W_\Lambda(x)$.

For example, in the canonical ensemble $W_\Lambda(x)$ is the total energy and the number of particles is fixed.

We can express the distribution of some other variable $U_{0\Lambda}(x)$ in this ensemble as before if we also include W among $U_0,\ldots, U_m$ and consider $Z_\Lambda(I_0 \times I \times J, a_0, a, b) =$

$$\sum_{x \in X_\Lambda} \exp[-a_0 U_{0\Lambda}(x) - a \cdot U_\Lambda(x) - b W_\Lambda(x)] = |\Lambda|^{-1}(U_{0\Lambda}(x), U_\Lambda(x), W_\Lambda(x)) I_0 \times I \times J$$

The probability distribution of $|\Lambda|^{-1} U_{0\Lambda}(x)$ is then given by

$$p_\Lambda(|\Lambda|^{-1} U_{0,\Lambda}(x) \in I_0; I, b) \equiv p_\Lambda(I_0; I, b) = \frac{Z_\Lambda(I_0 \times I \times R^1, 0, 0, b)}{Z_\Lambda(R^1 \times I \times R^1, 0, 0, b)} \equiv \frac{Z_\Lambda{}^b(I_0 \times I)}{Z_\Lambda{}^b(I)}$$

making the same abuse of notation as before.

As before, we have

$$\lim_{\Lambda \to \infty} |\Lambda|^{-1} \log Z_\Lambda(I_0 \times I \times J, a_0, a, b) = s(I_0 \times I \times J, a_0, a, b) = \sup_{\substack{u_0 \in I_0 \\ u \in I; w \in J}} (s(u_0, u, w) - a_0 u_0 - a \cdot u - b \cdot w)$$

with an extended entropy function $s(u_0, u, w)$ having the same properties as before and conjugate function $g(a_0 U_0 + a \cdot U + bW) \equiv g(a_0, a, b) = \sup_{u_0, u, w}(s(u_0, u, w) - a_0 u_0 - a \cdot u - bw)$.

For the limit of $Z_\Lambda{}^b(I_0 \times I)$ we then have

$$\begin{aligned} \lim_{\Lambda \to \infty} |\Lambda|^{-1} \log Z_\Lambda{}^b(I_0 \times I) &\equiv s^b(I_0 \times I) \\ &= s(I_0 \times I \times R^1, 0, 0, b) \\ &= \sup_{\substack{u_0 \in I_0 \\ u \in I; w \in R^1}} (s(u_0, u, w) - bw) \\ &= \sup_{u_0 \in I_0; u \in I} \sup_{w} (s(u_0, u, w) - bw) \\ &\equiv \sup_{u_0 \in I_0; u \in I} s^b(u_0, u) \end{aligned}$$

Hence for this ensemble we have the same situation as for the restricted ensemble with $s^b(u_0, u) = \sup_w(s(u_0, u, w) - bw)$ replacing $s(u_0, u)$ and $g^b(a_0, a) = g(a_0, a, b)$ replacing $g(a_0, a)$, because

$$g^b(a_0, a) = g(a_0, a, b) = \sup_{u_0, u}((\sup_w s(u_0, u, w) - bw) - a_0u_0 - a\cdot u)$$
$$= \sup_{u_0, u}(s^b(u_0, u) - a_0u_0 - a\cdot u)$$

It is easy to check that $s^b(u_0, u)$ is concave, u.s.c. as $s(u_0, u)$ before.[5]

The same discussion as for the restricted ensemble applied to the functions s^b and g^b hence gives the result corresponding to Theorem 1 for the mixed ensemble: There is a unique limit of the mixed state $\bar{p}_\Lambda(x; I, b)$ for any u corresponding to a in the duality between g^b and s^b if there are finitely many pure phases, $\{p_e, e = 1,..., f\}$, for $V = a\cdot U + bW$ and the equations

$$u_i = \sum_1^f (p_e, U_i)q_e, \quad i = 1,..., m; \qquad a = \sum_1^f q_e$$

determine the q_e uniquely.

This result shows in particular that if we compare the restricted ensemble determined by $u \in R^m$ corresponding to $a \in R^m$ having order parameters $U_1,..., U_n$ and the mixed ensemble determined by $(u_1,..., u_n)$ and $bW = \sum_{n+1}^m a_iU_i$, then we get precisely the same limits of the state, since in both cases we are the same point $V = a\cdot U = \sum_1^n a_iU_i + bW$ in B. That is, the restrictions not corresponding to the order parameters can be taken care of by a mixed ensemble correctly centered so that

$$bW = \sum_{n=1}^m a_iU_i.$$

Example. In the Ising model defined before we have the following situation when $a_1 = \beta > \beta_0$, $a_2 = 0$: $U_{1\Lambda}(x)$ = total interaction energy, $U_{2\Lambda}(x)$ = total magnetization. Since $U_{1\Lambda}(x)$ is even and $U_{2\Lambda}(x)$ is odd when $x \to -x$, we have

$$(p_\pm, U_1) = -g_1'(a) \qquad (p_\pm, U_2) = \pm m^*(\beta)$$

$[-m^*(\beta)$ is the right derivative $g_{2,+}'(a)]$. Hence $U_{2\Lambda}(x)$ is the order parameter, and the mixture is uniquely determined by the equations

$$u_2 = m^*(\beta)q_+ - m^*(\beta)q_-, \qquad 1 = q_+ + q_-$$

The remaining equation

$$u_1 = -g_1'(\beta, 0)$$

defines the correspondence between u and a.

Our results hence tell us that there is a unique limit of the microcanonical state for $u_1 = -g_1'(\beta, 0)$, $|u_2| \leqslant m^*(\beta)$, and that it is also the limit of the canonical state determined by β and u_2, $|u_2| \leqslant m^*(\beta)$.

In this model, the fact that in the canonical ensemble the system consists of a mixture of two phases occupying macroscopic regions having volume fractions $q_\pm$ has been proved in two dimensions and for low temperature in the detailed investigations by Minlos and Sinai.[7]

Remark. The problem of the equivalence of different ensembles in a finite container Λ has also been considered, e.g., by Thompson[10] and Dobrushin and Tirozzi.[1] They consider a restricted probability law defined on a "thin energy shell" $|\Lambda|^{-1}U_\Lambda(x) = u_\Lambda =$ fixed and $u_\Lambda \to u$ as $\Lambda \to \infty$. The analysis becomes more complicated in this case since one has to analyze the range of $U_\Lambda(x)$ as $x \in X_\Lambda$ more carefully; see Ref. 10, where equivalence of ensembles is obtained when there is no phase transition using "thermodynamic" probability estimates for the law of large numbers as done here. In Ref. 1 the equivalence between the canonical and grand canonical ensembles is considered for a lattice gas, i.e., $S = \{0, 1\}$ and $U_\Lambda(x)$ is equal to the number of particles in Λ. The method used is to estimate the conditional distribution of the state of a subsystem given that $U\Lambda(x) = u_\Lambda$ in the grand canonical ensemble using the local central limit theorem. This requires a more delicate analysis, and again only the situation when there is no phase transition is treated.

REFERENCES

1. R. L. Dobrushin and B. Tirozzi, The central limit theorem and the problem of equivalence of ensembles, *Comm. Math. Phys.* **54**:173 (1977).
2. R. Griffiths and D. Ruelle, Strict convexity ("continuity") of the pressure in lattice systems, *Comm. Math. Phys.* **23**:169 (1971).
3. O. Lanford, Entropy and equilibrium states in classical statistical mechanics, in *Statistical Mechanics and Mathematical Problems*, A. Lenard, ed. (Springer Lecture Notes in Physics 20).
4. J. Lebowitz and A. Martin-Löf, On the uniqueness of the equilibrium state for Ising spin systems, *Comm. Math. Phys.* **25**:276 (1972).
5. A. Martin-Löf, Lectures on statistical mechanics and the foundations of thermodynamics, Res. Rep. no. 105, Dept. of Math. Statistics, University of Stockholm (1977), to appear in *Springer Lecture Notes in Physics.*
6. A. Messger and S. Miracle-Solé, Equilibrium states of the two-dimensional Ising model in the two-phase region, *Comm. Math. Phys.* **40**:187–196 (1975).
7. R. A. Minlos and Ja. G. Sinai, The phenomenon of "phase separation" at low temperatures in some lattice models of a gas. I, *Math. USSR—Sbornik* **2**:335 (1967); II, *Trans. Moscow Math. Soc.* **19**:121 (1968).
8. R. T. Rockafellar, *Convex Analysis* (Princeton University Press, 1970).
9. D. Ruelle, *Statistical Mechanics* (Benjamin, 1969).
10. R. L. Thompson, Equilibrium states on thin energy shells, Mem. AMS no. 150 (1974).