Z. Wahrscheinlichkeitstheorie verw. Gebiete 32, 277–300 (1975)

Canonical Gibbs States, Their Relation to Gibbs States, and Applications to Two-Valued Markov Chains

Hans-Otto Georgii

0. Introduction

We consider a countable set S of *sites* each of which may be occupied by a particle or not. A *configuration* of particles on S is described by an element ω of the space $\Omega=\{0,1\}^S$ where $x\in S$ is meant to be occupied by a particle iff $\omega_x=1$. Ω is endowed with the σ-field $\mathscr{F}$ generated by the cylindric sets. A probability measure on $(\Omega,\mathscr{F})$ is called a *random field* or shorter a *state*. The local interaction structure of particles in equilibrium leads to the system of Gibbs distributions the latter being conditional probabilities of the type: given a configuration outside of a certain finite volume $V\subset S$, what are the probabilities of the configurations inside of V. A state whose local conditional probabilities are given by this prescribed system of nice versions is called *Gibbsian* [2, 12].

In the last few years, Gibbsian random fields have received a lot of attention so that they form a nice theory now, see [7, 14, 15, 18] for a survey. In particular, it is known that they are invariant measures (and under certain conditions all invariant measures) of certain Markovian interaction processes regulating the birth or death of a particle at $x\in S$ according to the aim of minimizing the interaction energy at x [11, e.g.]. More interesting for particle systems is another type of interaction process where the interaction induces jumps of the particles [17]. These processes which under natural conditions preserve the particle density have as natural candidates for the invariant measures a considerably larger class of states than the corresponding set of Gibbs states [10, see also 13]. This class may be characterized by the fact that its local conditional probabilities look like the Gibbsian ones only if in addition to the configuration outside of V the particle number in V is given. If a random field has this property we call it a *canonical Gibbs state*.

In this note we investigate the relation between Gibbs states and canonical Gibbs states. The basic technique is a comparison of variational principles. Our main result (Theorem (4.1)) may be understood as a generalization of the Finetti's theorem on exchangeable 0-1 variables [9] to interacting particle systems. Indeed, if there is no interaction then the canonical Gibbs states are nothing else than the symmetric states and the Gibbs states are product measures. From this point of view, our communication is a contribution to Spitzer's program of breaking away from the sphere of influence of stochastic independency [17].

What is the reason for the term "canonical"? The usual Gibbs states are characterized by an interaction potential describing the interaction of distinct particles and a socalled chemical potential which measures the likelihood for the particles to exist at all. This corresponds exactly to the grand canonical description of thermodynamic states. In the canonical thermodynamic ensemble the additional parameter to the interaction potential is the particle density. Our canonical Gibbs states are characterized by the interaction of distinct particles only. Hence if a random field has a particle density, its characterization by this density and the property of being a canonical Gibbs state is a canonical description.

The existence of particle densities necessitates some structure of the set S. Therefore, we shall restrict ourselves to the standard case when S is a d-dimensional cubic lattice, and we will concentrate most of our interest on the shift invariant states.

After the basic definitions in Section 1, we shall establish in Section 2 some fundamental results on the set of canonical Gibbs states. Using the Legendre relation between the canonical and the grand canonical free energies we find a variational characterization of all shift invariant canonical Gibbs states (Theorem (3.8)) and the characterization of the ergodic canonical Gibbs states as Gibbs states with respect to an appropriately chosen chemical potential (Theorem (4.1)). Section 4 contains a discussion of the consequences concerning phase transitions and the equivalence of ensembles.

In Section 5 we investigate one-dimensional systems with nearest neighbour interactions. We prove that the corresponding canonical Gibbs states have a delicate symmetry property and are necessarily shift invariant. The application of the preceding results yields an extension of the Finetti's theorem to certain Markov chains. In the final section we use our method of comparison of free energies for an investigation of the closed convex hull of all ergodic Markov states. This leads to a "microcanonical" description of states with a simple symmetry property via the particle density and the pair correlation. As a by-product we find that the ergodic Markov chains are most random in the class of all shift invariant states with the same expected particle density and expected pair correlation.

Finally, let us remark that the results of Sections 1 to 4 have a straightforward extension to the case where each site is allowed to carry more than one particle up to a finite maximal number.

Note. After this paper had been submitted for publication, I found out that R.L. Thompson has investigated the more general notion of states being defined by conditions on the energies with respect to certain potentials instead of conditions on the particle numbers: Equilibrium states on thin energy shells, Memoirs Amer. Math. Soc. 150 (1974). In the case of finite range potentials, Theorem (4.1) is contained in Thompson's theorem 3.1. The canonical Gibbs states, however, are much more easily and more completely treated by the methods below.

1. Definitions

As the set of sites we fix the d-dimensional lattice $S=\mathbb{Z}^d$, $d\geqq 1$. Then we may consider the *shift group* $\Theta=(\theta_x)_{x\in S}$ acting on $\Omega=\{0,1\}^S$ via

$$(\theta_x\omega)_y=\omega_{y-x} \qquad (x, y\in S).$$

This definition extends obviously to configurations that are defined only on a subset of S. Furthermore, we may shift any $V \subset S$ by $x \in S$ to $V+x$.

We denote by $\mathcal{S}$ the set of nonempty subsets V of S with finite cardinality $|V|$, by $\mathcal{F}_V$ the σ-field on Ω generated by the cylindric sets depending only on $V \subset S$, by X_V the projection of Ω onto $\Omega_V = \{0, 1\}^V$ $(V \subset S)$, $\omega_V = X_V(\omega)$, and for disjoint subsets V, W of S and $\omega \in \Omega_V$, $\zeta \in \Omega_W$ by $\omega\zeta \in \Omega_{V \cup W}$ the coupled configuration on $V \cup W$.

First we give the definitions leading to the notion of a Gibbs state.

(1.1) *Definition.* A *specification* Λ is a system

$$\left(\lambda_V(\omega|\eta)\right)_{\omega \in \Omega_V, \eta \in \Omega, V \in \mathcal{S}}$$

of probability kernels from $(\Omega, \mathcal{F}_{S \setminus V})$ to Ω_V (i.e., for each $\omega \in \Omega_V$, $\lambda_V(\omega|\cdot)$ is a $\mathcal{F}_{S \setminus V}$-measurable function on Ω, and for all $\eta \in \Omega$, $\lambda_V(\cdot|\eta)$ is a probability vector on Ω_V) with the properties

(S 1) *Consistency.* If $V \subset W \in \mathcal{S}$, $\omega \in \Omega_V$, $\eta \in \Omega$, then

$$\lambda_V(\omega|\eta) \sum_{\zeta \in \Omega_V} \lambda_W(\zeta \eta_{W \setminus V}|\eta) = \lambda_W(\omega \eta_{W \setminus V}|\eta).$$

(S 2) *Continuity.* For any $V \in \mathcal{S}$, $\omega \in \Omega_V$, the function $\lambda_V(\omega|\cdot)$ is continuous with respect to the product topology on Ω.

(S 3) *Shift Invariance.* For all $V \in \mathcal{S}$, $\omega \in \Omega_V$, $\eta \in \Omega$, $x \in S$

$$\lambda_V(\omega|\eta) = \lambda_{V+x}(\theta_x \omega|\theta_x \eta).$$

All relevant specifications Λ can be represented by an interaction potential (see [7, 14], e.g.) in the sense of

(1.2) *Definition.* (a) An *interaction potential* is a map

$$\Phi: \mathcal{S} \to \mathbb{R}$$

with the properties

(P 1) $$\|\Phi\| := \sum_{0 \in A \in \mathcal{S}} |\Phi(A)| < \infty,$$

(P 2) $$\Phi(A) = \Phi(A+x) \quad (A \in \mathcal{S}, x \in S).$$

(b) The *energy* of $\omega \in \Omega$ in $V \in \mathcal{S}$ under Φ is defined by

$$E_V(\omega) = \sum_{A \in \mathcal{S}: A \cap V \neq \emptyset} \Phi(A) \omega^A,$$

where $\omega^A = \prod_{x \in A} \omega_x$.

(c) By Λ_Φ we denote the *Gibbsian specification* with respect to Φ defined by

$$\lambda_V(\omega|\eta) = Z_V(\eta)^{-1} \exp[-E_V(\omega \eta_{S \setminus V})].$$

The normalizing constant $Z_V(\eta)$ is called the *partition function.* It is easily verified that Λ_Φ is a specifiction, indeed [5, 7].

Now we consider states on $(\Omega, \mathcal{F})$. Denote by $\mathfrak{S}$ the set of states and by $\mathfrak{S}_\theta$ the set of shift invariant states, i.e., of all states μ such that

$$\mu(\theta_x A) = \mu(A) \quad (A \in \mathcal{F}, x \in S).$$

(1.3) *Definition.* One says that $\mu \in \mathfrak{S}$ is specified by Λ, and μ is called a *Gibbs state*, if for all $V \in \mathscr{S}$, $\omega \in \Omega_V$

$$\mu[X_V = \omega \mid \mathscr{F}_{S \setminus V}] = \lambda_V(\omega \mid \cdot) \qquad \mu\text{-a.s.}$$

The basic theorem by Dobrushin [2] states that *for all specifications Λ the set $\mathfrak{G}(\Lambda)$ of all Gibbs states specified by Λ and the set $\mathfrak{G}_\theta(\Lambda) = \mathfrak{G}(\Lambda) \cap \mathfrak{S}_\theta$ are nonempty, convex, and weakly compact.* Furthermore, *they are simplices in the sense of Choquet, and their extreme points are characterized by the property of being trivial on the tail field*

$$\mathscr{F}_\infty = \bigcap_{V \in \mathscr{S}} \mathscr{F}_{S \setminus V}$$

and on the σ-field $\mathscr{I}$ of shift invariant sets, respectively [4, 5, 7, 12, 14, 15].

In complete analogy to the above we now introduce the concept of canonical Gibbs states. Denote by $\Omega_{V,N}$ the set of all $\omega \in \Omega_V$ whose particle number

$$N(\omega) = \sum_{x \in V} \omega_x = |\{x \in V : \omega_x = 1\}|$$

is given by N, $0 \leqq N \leqq |V|$. Sometimes we shall use the same symbol for the set $\{N(X_V) = N\}$.

(1.4) *Definition.* A *canonical specification Γ* is a system

$$\big(\gamma_{V,N}(\omega \mid \eta)\big)_{\omega \in \Omega_V, \eta \in \Omega, 0 \leqq N \leqq |V|, V \in \mathscr{S}}$$

of probability kernels from $(\Omega, \mathscr{F}_{S \setminus V})$ to Ω_V with the properties

(CS 0) $\gamma_{V,N}(\Omega_{V,N} \mid \eta) = 1$ for all $\eta \in \Omega$,

(CS 1) If $V \subset W \in \mathscr{S}$, $\omega \in \Omega_W$, $\eta \in \Omega$, $0 \leqq N \leqq |W|$, then

$$\gamma_{W,N}(\omega \mid \eta) = \gamma_{V,N(\omega_V)}(\omega_V \mid \omega \eta_{S \setminus W}) \sum_{\zeta \in \Omega_{V,N(\omega_V)}} \gamma_{W,N}(\zeta \omega_{W \setminus V} \mid \eta)$$

i.e., $\gamma_{V,n}(\cdot \mid \eta)$ is the conditional probability on Ω_V with respect to the measure $\gamma_{W,N}(\cdot \mid \eta)$ under the condition

$$\{N(X_V) = n, X_{W \setminus V} = \eta_{W \setminus V}\} \qquad \text{if } n + N(\eta_{W \setminus V}) = N.$$

(CS 2) For all $V \in \mathscr{S}$, $0 \leqq N \leqq |V|$, $\omega \in \Omega_{V,N}$ the function $\gamma_{V,N}(\omega \mid \cdot)$ is continuous.

(CS 3) If $V \in \mathscr{S}$, $\omega \in \Omega_V$, $\eta \in \Omega$, $0 \leqq N \leqq |V|$, $x \in S$, then

$$\gamma_{V,N}(\omega \mid \eta) = \gamma_{V+x,N}(\theta_x \omega \mid \theta_x \eta).$$

(1.5) *Remark.* If a specification Λ has the property

$$\lambda_V(\Omega_{V,N} \mid \cdot) > 0 \qquad \text{for all } V \in \mathscr{S}, 0 \leqq N \leqq |V|,$$

one obtains a canonical specification Γ_Λ by setting

$$\gamma_{V,N}(\cdot \mid \eta) = \lambda_V(\cdot / \Omega_{V,N} \mid \eta) \qquad (V \in \mathscr{S}, \eta \in \Omega, 0 \leqq N \leqq |V|).$$

Proof. The properties (CS 0), (CS 2), and (CS 3) are evident. To show (CS 1) it is sufficient to consider the case $N(\omega) = N$. Then (CS 1) follows easily from (*S* 1). □

Mainly we are interested in *canonical Gibbs specifications* Γ_Φ with respect to some potential Φ obtained from Λ_Φ as in (1.5). Γ_Φ is given by

$$(1.6)\qquad \gamma_{V,N}(\omega|\eta)=\begin{cases} Z_{V,N}(\eta)^{-1}\exp\left[-E_V(\omega\eta_{S\setminus V})\right] & \text{if } \omega\in\Omega_{V,N} \\ 0 & \text{otherwise.}\end{cases}$$

The normalizing constant $Z_{V,N}(\eta)$ is called the *canonical partition function*, and $\gamma_{V,N}(\cdot|\eta)$ is the *canonical Gibbs distribution on* $\Omega_{V,N}$ *under the boundary condition* η.

(1.7) *Remark.* In (1.6) the value $\varphi := -\Phi(A)(|A|=1)$ cancels out and therefore has no effect if we consider canonical Gibbs specifications. Therefore, in the canonical context we make the convention that $\varphi=0$. We call such a potential a *canonical potential.* Each potential can be uniquely represented by a canonical potential and its so-called *chemical potential* φ.

Consider now the σ-field $\mathscr{G}_V$ $(V\in\mathscr{S})$ of all sets of the form

$$\{N(X_V)=N\}\cap A \qquad (A\in\mathscr{F}_{S\setminus V},\, 0\leqq N\leqq|V|).$$

Note that for any canonical specification Γ the function

$$\gamma_V(\omega|\cdot) := \gamma_{V,N(X_V)}(\omega|\cdot) = 1_{\Omega_{V,N(\omega)}}\,\gamma_{V,N(\omega)}(\omega|\cdot)$$

is $\mathscr{G}_V$-measurable, and for all $\eta\in\Omega$, $\gamma_V(\cdot|\eta)$ is a probability vector on Ω_V. Hence $\gamma_V(\cdot|\cdot)$ plays the same role for the σ-field $\mathscr{G}_V$ as $\lambda_V(\cdot|\cdot)$ for the σ-field $\mathscr{F}_{S\setminus V}$. This leads to the definition

(1.8) *Definition.* $\mu\in\mathfrak{S}$ is said to be *canonically specified* by Γ and is called a *canonical Gibbs state* for Γ if for any $V\in\mathscr{S}$, $\omega\in\Omega_V$

$$\mu[X_V=\omega|\mathscr{G}_V]=\gamma_V(\omega|\cdot) \qquad \mu\text{-a.s.},$$

that is, if for all $A\in\mathscr{F}_{S\setminus V}$ the equality

$$(1.9)\qquad \int_A d\mu\, 1_{[X_V=\omega]} = \int_{A\cap\Omega_{V,N(\omega)}} d\mu\,\gamma_{V,N(\omega)}(\omega|\cdot)$$

is true. $\mathfrak{C}=\mathfrak{C}(\Gamma)$ denotes the set of all canonical Gibbs states for Γ, and $\mathfrak{C}_\theta = \mathfrak{C}_\theta(\Gamma) := \mathfrak{C}(\Gamma)\cap\mathfrak{S}_\theta$.

(1.10) *Remark.* In the situation of (1.5),

$$\mathfrak{G}(\Lambda)\subset\mathfrak{C}(\Gamma_\Lambda).$$

Proof. For all $V\in\mathscr{S}$, $\omega\in\Omega_V$, $A\in\mathscr{F}_{S\setminus V}$, $\mu\in\mathfrak{G}(\Lambda)$ we have

$$\begin{aligned}\int_A 1_{[X_V=\omega]}\,d\mu &= \int_A \lambda_V(\omega|\cdot)\,d\mu = \int_A \gamma_{V,N(\omega)}(\omega|\cdot)\,\lambda_V(\Omega_{V,N(\omega)}|\cdot)\,d\mu \\ &= \int_A E_\mu[\gamma_{V,N(\omega)}(\omega|\cdot)\,1_{\Omega_{V,N(\omega)}}|\mathscr{F}_{S\setminus V}]\,d\mu \\ &= \int_A 1_{\Omega_{V,N(\omega)}}\,\gamma_{V,N(\omega)}(\omega|\cdot)\,d\mu. \quad \square\end{aligned}$$

Consequently, if $\Gamma=\Gamma_\Lambda$ for some Λ then the existence theorem for Gibbs states implies that $\mathfrak{C}_\theta(\Gamma)$ is non-empty. But the existence problem for canonical Gibbs states is by far easier. Consider the unit masses ε_0 and ε_1 on the constant configurations $\mathbb{0}$ and $\mathbb{1}$, respectively, where $\mathbb{0}_x=0$ $(x\in S)$ and $\mathbb{1}_x=1$ $(x\in S)$.

(1.11) *Remark.* If a system Γ has property (CS 0), the random fields ε_0 and ε_1 belong to $\mathfrak{C}_\theta(\Gamma)$.

Proof. (1.9) is verified for ε_0, e.g., by observing that $|\Omega_{V,0}|=1$, and therefore $\gamma_{V,0}(\mathbb{0}_V|\cdot)=1$. □

2. The Simplices $\mathfrak{C}$ and $\mathfrak{C}_\theta$

We fix a canonical specification Γ and consider $\mathfrak{C}=\mathfrak{C}(\Gamma)$. First we observe that canonical specifications fit into the most general setting of specified random fields which has been considered by Föllmer [4] and, in the sequel, by Preston [15, Section 10].

Define probability kernels $\pi_V(\cdot, A)$ from $(\Omega, \mathscr{G}_V)$ to $(\Omega, \mathscr{F})$ by

$$\pi_V(\eta, [X_V=\omega]\cap A)=1_A(\eta)\,\gamma_V(\omega|\eta)=1_{A\cap\Omega_{V,N}}(\eta)\,\gamma_{V,N}(\omega|\eta)$$

where $V\in\mathscr{S}$, $\eta\in\Omega$, $A\in\mathscr{F}_{S\setminus V}$, and $\omega\in\Omega_{V,N}$.

(2.1) **Lemma.** *The kernels π_V $(V\in\mathscr{S})$ have the following properties:*

(a) *If $V\subset W\in\mathscr{S}$ then $\pi_W\pi_V=\pi_W$.*
(b) *If f is a continuous function on Ω, then $\pi_V f$ is continuous as well.*
(c) *$\mu\in\mathfrak{C}$ if and only if $\mu\pi_V=\mu$ for any $V\in\mathscr{S}$.*

Proof. (a) and (b) are easily derived from (CS 1) and (CS 2) and (c) from (1.9). □

Furthermore, observe that for all $V\in\mathscr{S}$ the σ-field $\mathscr{G}_V$ consists of all events in $\mathscr{F}$ which are invariant under permutations of the sites in V. Accordingly,

(2.2) *Remark.* If $V\subset W\in\mathscr{S}$ then $\mathscr{G}_V\supset\mathscr{G}_W$.

The observations (2.1) and (2.2) show that we are exactly in the situation of [4] and [15, Section 10]. From there we deduce a collection of results which we list now.

(2.3) **Theorem** (from [15]). *The sets $\mathfrak{C}$ and $\mathfrak{C}_\theta$ are convex and weakly compact. More precisely, they are simplices in the sense of Choquet.*

The next theorem states that $\mathfrak{C}$ coincides with a certain set $\mathfrak{C}_\infty$ which has been introduced in [10] as the set of equilibrium states for particle jump processes. $\mathfrak{C}_\infty$ is defined as the set of all weak limits of states μ^k whose marginal distributions in Ω_{V_k} are of the form

$$\mu^k[X_{V_k}=\omega]=\int d\nu^k\,\gamma_{V_k}(\omega|\cdot)$$

with certain probability measures ν^k and a sequence $(V_k)_k$ in $\mathscr{S}$ increasing to S.

(2.4) **Theorem** (from [15]). $\mathfrak{C}=\mathfrak{C}_\infty$.

For the characterization of the extremal points of $\mathfrak{C}$ the σ-field

$$\mathscr{G}_\infty=\bigcap_{V\in\mathscr{S}}\mathscr{G}_V$$

plays an essential role. $\mathscr{G}_\infty$ contains the tail field $\mathscr{F}_\infty$ and the information on the asymptotic behaviour of the particle number in large volumes. We denote by ex K the set of extreme points of a convex set K.

(2.5) **Theorem** (from [4, 15]). (a) *$\mu\in\mathfrak{C}$ is extremal in $\mathfrak{C}$ if and only if $\mu(A)=0$ or 1 for all $A\in\mathscr{G}_\infty$.*

(b) *Let $\mu, \nu\in\mathfrak{C}$. Then $\mu=\nu$ if and only if $\mu=\nu$ on $\mathscr{G}_\infty$.*

(c) *Distinct extreme points of $\mathfrak{C}$ are mutually singular.*

Statement (a) shows that we cannot expect that ex $\mathfrak{G}(\Lambda)$ is a subset of ex $\mathfrak{C}(\Gamma_\Lambda)$ whenever we are in the situation of (1.10).

We have to distinguish between nice extreme points of $\mathfrak{C}$ ("phases") and exceptional extreme points which we cannot exclude in general.

(2.6) **Corollary.** *If $\mu\in$ ex $\mathfrak{C}$, then for any sequence $V\uparrow S$ either*

$$\mu\big[\lim_{V\uparrow S} N(X_V)/|V| \text{ exists}\big]=0,$$

or there is some $\rho\in[0, 1]$ such that

$$\mu\big[\lim_{V\uparrow S} N(X_V)/|V|=\rho\big]=1.$$

Proof. The events $\{\limsup N(X_V)/|V|>\alpha\}$ and $\{\liminf N(X_V)/|V|<\beta\}$ belong to $\mathscr{G}_\infty$, hence they have μ-measure 0 or 1. The result follows by letting α and β vary over the rational numbers. □

Furthermore, we see that a phase with particle density ρ can be approximated by canonical Gibbs distributions in finite volumes V and a particle number nearly equal to $\rho|V|$. As in [6, 4, 15] one proves

(2.7) **Corollary.** *Each $\mu\in$ ex $\mathfrak{C}$ is the weak limit of states μ^V with marginal distributions $\gamma_{V,N(\eta_V)}(\cdot|\eta)$ in Ω_V where $V\uparrow S$. η can be chosen from a set of μ-measure 1.*

From [4] we obtain the following explicit integral representation of any $\mu\in\mathfrak{C}$. Fix some sequence $V\uparrow S$. Then for μ-a.a. $\eta\in\Omega$ the weak limit ν^η of $\gamma_{V,N(\eta_V)}(\cdot|\eta)$ (more precisely, of $\pi_V(\eta,\cdot)$) exists, belongs to ex $\mathfrak{C}$, and

$$\mu=\int \nu^\eta\, \mu|_{\mathscr{G}_\infty}(d\eta). \tag{2.8}$$

Now we consider $\mathfrak{C}_\theta$ and the σ-field $\mathscr{I}$ of shift invariant sets. Then an analogue of (2.8), with $\mathscr{G}_\infty$ replaced by $\mathscr{I}$, is true [4] whence we obtain

(2.9) **Theorem.** *$\mu\in\mathfrak{C}_\theta$ is extremal in $\mathfrak{C}_\theta$, if and only if μ is ergodic (i.e., $\mu(A)=0$ or 1 for $A\in\mathscr{I}$).*

This can be proved, too, by using the fact that for a shift invariant state μ the σ-field $\mathscr{I}$ is a.s. contained in $\mathscr{F}_\infty$ (see [6] for a proof, e.g.) and hence in $\mathscr{G}_\infty$ and the trivial fact that if $\mu\in\mathfrak{C}$, h is $\mathscr{G}_\infty$-measurable, and $h\mu\in\mathfrak{S}$ then $h\mu\in\mathfrak{C}$.

Using Hunt's extension of the martingale convergence theorem one derives from (2.9) as in [6, 4]:

(2.10) **Corollary.** *Each $\mu\in$ ex $\mathfrak{C}_\theta$ is the weak limit of states μ^V whose marginal distributions in Ω_V are given by*

$$\mu^V[X_V=\omega]=|V|^{-1}\sum_{x\in V}\gamma_{V+V+x,N(\eta_{V+V})}(X_V=\omega|\theta_x\eta)$$

where V runs through a sequence of cubes to S. η can be chosen from a set of μ-measure 1.

Notice that $N(\eta_{V+V})/|V+V|$ has an a.s. constant limit ρ. Since in the situation of (1.10) $\operatorname{ex}\mathfrak{G}_\theta(\Lambda)\subset\operatorname{ex}\mathfrak{C}_\theta(\Gamma_\Lambda)$, (2.10) provides some information in addition to the fact that each $\mu\in\operatorname{ex}\mathfrak{G}_\theta(\Lambda)$ can be approximated by

$$|V|^{-1}\sum_{x\in V}\lambda_{V+V}(X_V=\omega\,|\,\theta_x\eta)$$

as proved in [6].

We leave it to the interested reader to give a direct proof of the theorems in this section by a slight modification of the methods in [7, 14].

3. A Variational Characterization of $\mathfrak{C}_\theta$

Let us fix a canonical potential Φ and consider the set $\mathfrak{C}_\theta(\Phi)=\mathfrak{C}_\theta(\Gamma_\Phi)$. We are going to define a certain function $c(\cdot,\Phi)$ on $\mathfrak{S}_\theta$ vanishing exactly on $\mathfrak{C}_\theta(\Phi)$. c is defined as a specific quantity per volume, i.e., by a limiting process $V\uparrow S$ where we have to require that V increases in a nice way. Hence, we fix some sequence of "cubes" increasing to S, and throughout this section we mean by "$V\uparrow S$" that V runs through this sequence.

First let us recall what is known about the variational characterization of $\mathfrak{G}_\theta(\Psi)$ for any potential Ψ. If μ is shift invariant, the *specific entropy*

$$s(\mu)=\lim_{V\uparrow S}|V|^{-1}\int\mu(d\omega)\log\mu[X_V=\omega_V] \tag{3.1}$$

exists [3, 5, 7, e.g.]. $s(\cdot)$ is an affine upper semicontinuous function on $\mathfrak{S}_\theta$, for a proof see [7], e.g. The *specific energy* of μ with respect to Ψ is defined by

$$e(\mu,\Psi)=\int\mu(d\omega)\sum_{0\in A\in\mathscr{S}}\frac{\Psi(A)}{|A|}\,\omega^A$$

and the *specific free energy* of μ with respect to Ψ by

$$f(\mu,\Psi)=e(\mu,\Psi)-s(\mu).$$

The theorem of Lanford and Ruelle [12, 3] states that *any $\mu\in\mathfrak{S}_\theta$ belongs to $\mathfrak{G}_\theta(\Psi)$ if and only if $f(\cdot,\Psi)$ attains its minimal value $-P(\Psi)$ at μ.* $P(\Psi)$ is called the *pressure* or the *specific free Gibbs energy* and is obtained as the limit of the logarithms of the partition functions for Ψ:

$$P(\Psi)=\lim_{V\uparrow S}|V|^{-1}\log Z_V(\eta;\Psi)$$

where η runs through an arbitrary sequence in Ω.

Now let us consider the corresponding canonical quantities. First we need the existence of the *Helmholtz free energy per volume* $g(\rho)=g(\rho,\Phi)$ first proved by van Hove and some properties due to Ruelle, Fisher, Dobrushin, and Minlos.

(3.2) **Lemma.** *If $V\uparrow S$ and N tends to infinity such that $N/|V|$ has a limit $\rho\in[0,1]$, then for any choice of a sequence of boundary conditions $\zeta\in\Omega$ the limit*

$$g(\rho,\Phi)=\lim|V|^{-1}\log Z_{V,N}(\zeta,\Phi)$$

exists. $g(\cdot, \Phi)$ is concave and has a continuous derivative decreasing from ∞ to $-\infty$, if ρ increases from 0 to 1. Furthermore, for any $\varphi \in \mathbb{R}$ the Legendre relation

$$P(\Phi, \varphi) = \max_{0 \leqq \rho \leqq 1} [g(\rho, \Phi) + \varphi \rho] \tag{3.3}$$

holds.

Proof. If Φ is a pair potential (i.e., $\Phi(A)=0$ if $|A|>2$) and $\zeta = \mathbb{0}$, a complete proof is given in [5]. This proof can be extended in a straightforward way to general Φ. The only thing that has to be replaced is the symmetry argument in [5, Section 2.1] leading from the uniform convergence of the functions $g_V(\rho)$ (see [5, Section 1.1] for the definition) on each interval $[0, \alpha]$ ($\alpha < 1$) to the uniform convergence on $[0, 1]$. Obviously, this can be achieved also by showing that for all $\varepsilon > 0$ there is an $\alpha < 1$ such that for sufficiently large V and $\alpha |V| \leqq N \leqq |V|$ the inequality

$$\big| |V|^{-1} \log Z_{V,N}(\mathbb{0}) - |V|^{-1} \log Z_{V,|V|}(\mathbb{0}) \big| < \varepsilon$$

holds. To show this observe that

$$|V|^{-1} \log \frac{Z_{V,N}(\mathbb{0})}{Z_{V,|V|}(\mathbb{0})} = |V|^{-1} \log \sum_{\omega \in \Omega_{V,N}} \exp\Big[\sum_{A \subset V} \Phi(A) - \sum_{A \subset \{x:\, \omega_x = 1\}} \Phi(A) \Big].$$

For the terms under the summation use the bound

$$\Big| \sum_{\substack{A \subset V \\ A \setminus \{x:\, \omega_x = 1\} \neq \emptyset}} \Phi(A) \Big| \leqq (|V| - N) \|\Phi\| \leqq |V| (1-\alpha) \|\Phi\| \qquad (\omega \in \Omega_{V,N},\ N \geqq \alpha |V|),$$

and the cardinality of $\Omega_{V,N}$ can be handled by Stirling's formula showing that $|V|^{-1} \log |\Omega_{V,N}|$ has a limit smaller than

$$-\alpha \log \alpha - (1-\alpha) \log (1-\alpha)$$

if $|V| \to \infty$ and $N \geqq \alpha |V|$, $\alpha > \frac{1}{2}$. This implies the uniform convergence of $g_V(\rho)$ on $[0, 1]$ to the continuous extension of $g(\rho)$ from $[0, 1[$ to $[0, 1]$. This is all that is needed for the proof of (3.3) along the lines of [5, Section 3.2].

Finally we have to remark that the limit $g(\rho)$ does not depend on the special choice of the boundary condition $\zeta = \mathbb{0}$. Indeed, for any $\zeta \in \Omega$ we have

$$\left| \log \frac{Z_{V,N}(\zeta)}{Z_{V,N}(\mathbb{0})} \right| \leqq \Delta(V) := \sum_{\substack{A \cap V \neq \emptyset \\ A \setminus V \neq \emptyset}} |\Phi(A)|,$$

and it is known that $\lim_{V \uparrow S} |V|^{-1} \Delta(V) = 0$ (for a proof see [3], e.g.). □

Consider now the particle density

$$\rho(\omega) = \limsup_{V \uparrow S} N(\omega_V)/|V|$$

of a configuration $\omega \in \Omega$. If $\mu \in \mathfrak{S}_\theta$ then the d-dimensional ergodic theorem (cf. [5], e.g.) asserts that for a.a. $\omega \in \Omega$ $\rho(\omega)$ is in fact a limit and defines a version of $E_\mu[X_0 | \mathscr{I}](\omega)$.

Now we are in the position to define for any $\mu \in \mathfrak{S}_\theta$ the *excess canonical free energy of* μ *with respect to* Φ *by*

$$c(\mu, \Phi) = e(\mu, \Phi) - s(\mu) - E_\mu[g(\rho(\cdot), \Phi)].$$

Note that Φ is a canonical potential so that there is no one-particle term in $e(\mu, \Phi)$.

The following remark implies that the function $c(\cdot, \Phi)$ attains its minimum on a convex compact subset of $\mathfrak{S}_\theta$. We shall see below that this set coincides with $\mathfrak{C}_\theta(\Phi)$, indeed.

(3.4) *Remark.* The function $c(\cdot, \Phi)$ on $\mathfrak{S}_\theta$ is affine and lower semicontinuous.

Proof. Clearly, $c(\cdot, \Phi)$ is affine. $e(\cdot, \Phi)$ is the expectation of a continuous function and thereby continuous. The entropy is upper semicontinuous. In order to show that $E_\mu[g(\rho(\cdot), \Phi)]$ is a lower semicontinuous function of μ consider the cubes $V_n = [0, 2^n - 1]^d \in \mathscr{S}$ and the continuous functions

$$\rho_n(\omega) = |V_n|^{-1} \sum_{x \in V_n} \omega_x$$

on Ω. ρ_{n+1} is the arithmetic mean of the functions $\rho_n \circ \theta_x$ where the d coordinates of x are equal to 0 or 2^n. Using the concavity of $g(\cdot, \Phi)$ we see that for any shift invariant μ

$$E_\mu[g(\rho_{n+1}(\cdot), \Phi)] \geqq E_\mu[g(\rho_n(\cdot), \Phi)]$$

showing that the function $\mu \to E_\mu[g(\rho(\cdot), \Phi)]$ is the pointwise limit of an increasing sequence of continuous functions. □

(3.5) **Lemma.** *For any* $\mu \in \mathfrak{S}_\theta$

$$c(\mu, \Phi) = \lim_{V \uparrow S} |V|^{-1} \int \mu(d\omega) \log \frac{\mu[X_V = \omega_V | N(X_V) = N(\omega_V)]}{\gamma_{V, N(\omega_V)}(\omega_V | \zeta)},$$

where ζ *runs through an arbitrary sequence in* Ω.

Proof. Let us write

$$|V|^{-1} \int \mu(d\omega) \log \frac{\mu[X_V = \omega_V | N(X_V) = N(\omega_V)]}{\gamma_{V, N(\omega_V)}(\omega_V | \zeta)}$$

$$= |V|^{-1} \int \mu(d\omega) \log \mu[X_V = \omega_V] - |V|^{-1} \int \mu(d\omega) \log(\Omega_{V, N(\omega_V)})$$
$$+ |V|^{-1} \int \mu(d\omega) E_V(\omega_V \zeta_{S \setminus V}) + \int \mu(d\omega) |V|^{-1} \log Z_{V, N(\omega_V)}(\zeta).$$

If $V \uparrow S$ the first term tends to $-s(\mu)$. The second term can be written as an entropy

$$|V|^{-1} \sum_{N=0}^{|V|} \mu(\Omega_{V,N}) \log \mu(\Omega_{V,N})$$

being bounded by $|V|^{-1} \log(|V|+1)$ and thereby tending to zero. The third term converges to $e(\mu, \Phi)$, see [3] for a proof. Lemma (3.2) guaranties that the integrand in the fourth term converges a.s. to $g(\rho(\omega), \Phi)$. Hence we have only to show that the functions $|V|^{-1} \log Z_{V, N(X_V)}(\zeta)$ are uniformly bounded. But obviously

$$e^{-|V| \, \|\Phi\|} \leqq Z_{V,N}(\zeta) = \sum_{\omega \in \Omega_{V,N}} \exp[-E_V(\omega \zeta_{S \setminus V})] \leqq 2^{|V|} e^{|V| \, \|\Phi\|}. \quad \square$$

(3.6) **Proposition.** *For all* $\mu \in \mathfrak{S}_\theta$, $c(\mu, \Phi) \geqq 0$. *If* $\mu \in \mathfrak{C}_\theta(\Phi)$ *then* $c(\mu, \Phi) = 0$.

Proof. For any $\omega \in \Omega_V$, $\zeta, \eta \in \Omega$ we have

$$|E_V(\omega \eta_{S \setminus V}) - E_V(\omega \zeta_{S \setminus V})| \leqq \Delta(V)$$

and therefore

$$e^{-2\Delta(V)} \leqq \gamma_{V, N(\omega)}(\omega|\eta)/\gamma_{V, N(\omega)}(\omega|\zeta) \leqq e^{2\Delta(V)}.$$

Furthermore, if we wish to estimate

$$\int \mu(d\omega) \log \frac{\mu[X_V = \omega_V | N(X_V) = N(\omega_V)]}{\gamma_{V, N(\omega_V)}(\omega_V|\zeta)}$$

for some $\mu \in \mathfrak{C}_\theta(\Phi)$ it is sufficient to consider only those $\omega \in \Omega$ for which

$$\mu[N(X_V) = N(\omega_V)] > 0.$$

For these ω we may write

$$\frac{\mu[X_V = \omega_V | N(X_V) = N(\omega_V)]}{\gamma_{V, N(\omega_V)}(\omega_V|\zeta)} = \int \mu(d\eta | \Omega_{V, N(\omega_V)}) \frac{\gamma_{V, N(\omega_V)}(\omega_V|\eta)}{\gamma_{V, N(\omega_V)}(\omega_V|\zeta)}$$

and obtain

$$\left|\log \mu[X_V = \omega_V | N(X_V) = N(\omega_V)]/\gamma_{V, N(\omega_V)}(\omega_V|\zeta)\right| \leqq 2\Delta(V).$$

Now apply Lemma (3.5) and the fact that $\lim_{V \uparrow S} \Delta(V)/|V| = 0$ to deduce that $c(\mu, \Phi) = 0$ for $\mu \in \mathfrak{C}_\theta(\Phi)$.

To show that $c(\cdot, \Phi)$ is nonnegative write its approximation term introduced in (3.5) as

$$|V|^{-1} \sum_{N=0}^{|V|} \mu(\Omega_{V,N}) \sum_{\omega \in \Omega_{V,N}} \mu[X_V = \omega | N(X_V) = N] \log \frac{\mu[X_V = \omega | N(X_V) = N]}{\gamma_{V,N}(\omega|\zeta)}.$$

The inner sum has the form of a relative entropy which is known to be non-negative [3, 7]. □

(3.7) **Theorem.** *Let* $\mu \in \mathfrak{S}_\theta$ *and* $-\infty \leqq \varphi \leqq +\infty$. *Then the following statements are equivalent:*

(a) $\mu \in \mathfrak{G}_\theta(\Phi, \varphi)$

(b) $c(\mu, \Phi) = 0$ *and* $\dfrac{\partial}{\partial \rho} g(\rho(\cdot), \Phi) = -\varphi$ *μ-a.s.*

Proof. First we consider the case where the chemical potential φ is real. Assume that $\mu \in \mathfrak{G}_\theta(\Phi, \varphi)$. Then from (1.10) and (3.6) we obtain that $c(\mu, \Phi) = 0$. The second assertion is known [5], but let us give another argument. Observe that

$$\begin{aligned} 0 &= f(\mu; \Phi, \varphi) + P(\Phi, \varphi) - c(\mu, \Phi) \\ &= E_\mu[P(\Phi, \varphi) - \varphi\rho(\cdot) - g(\rho(\cdot), \Phi)] \end{aligned}$$

and apply (3.3) to see that

$$g(\rho(\cdot), \Phi) + \varphi\rho(\cdot) = \max_{0 \leqq \rho \leqq 1} [g(\rho, \Phi) + \varphi\rho] \quad \text{a.s.}$$

This implies the second statement. Conversely, if (b) is true then the same argument shows that

$$f(\mu;\Phi,\varphi)=-P(\Phi,\varphi)$$

proving (a) according to Lanford/Ruelle's variational principle. If $\varphi=-\infty$ then $\Lambda_{(\Phi,\varphi)}$ is given by

$$\lambda_V(\omega|\cdot)=\begin{cases}1 & \text{if } N(\omega)=0\\ 0 & \text{otherwise}\end{cases}$$

so that $\mathfrak{G}(\Phi,\varphi)=\{\varepsilon_0\}$. On the other hand $\frac{\partial}{\partial\rho}g(\rho,\Phi)=+\infty$ if and only if $\rho=0$. Now the statement follows from the fact that ε_0 is the unique shift invariant state with particle density 0. Similarly, the case $\varphi=+\infty$ leads to the unit mass ε_1. □

(3.8) **Theorem.** *Let μ be any shift invariant state. Then $c(\mu,\Phi)=0$ if and only if $\mu\in\mathfrak{C}_\theta(\Phi)$.*

Proof. The "if" part is done in (3.6). Suppose now that $c(\mu,\Phi)=0$. We represent μ as the gravicenter of a probability measure Q^μ on the ergodic states:

$$\mu=\int_{\mathrm{ex}\,\mathfrak{S}_\theta}\nu\,Q^\mu(d\nu).$$

Since $c(\cdot,\Phi)$ is affine and semicontinuous we conclude that

$$0=c(\mu,\Phi)=\int_{\mathrm{ex}\,\mathfrak{S}_\theta}c(\nu,\Phi)\,Q^\mu(d\nu)$$

and thereby $c(\nu,\Phi)=0$ for Q^μ-a.a. ν.

Furthermore, for any ergodic ν the particle density $\rho(\cdot)$ is a.s. constant. Hence the preceding theorem guaranties that Q^μ-a.a. ν belong to the subset

$$\bigcup_{-\infty\leqq\varphi\leqq\infty}\mathrm{ex}\,\mathfrak{G}_\theta(\Phi,\varphi)\quad\text{of } \mathrm{ex}\,\mathfrak{C}_\theta(\Phi).$$

Thus $\mu\in\mathfrak{C}(\Phi)$. □

In a more special situation the "only if" part can also be deduced from [10].

Note that we have proved the relation

$$\mathrm{ex}\,\mathfrak{C}_\theta(\Phi)\subset\mathrm{ex}\,\mathfrak{S}_\theta$$

a second time, and furthermore, that the representing measure of any $\mu\in\mathfrak{C}_\theta(\Phi)$ on the ergodic states is carried by $\mathrm{ex}\,\mathfrak{C}_\theta(\Phi)$. Accordingly, $\mathfrak{C}_\theta(\Phi)$ is a simplex since $\mathfrak{S}_\theta$ is a simplex.

Jensen's inequality combined with (3.7) and (3.8) or the Lanford and Ruelle theorem combined with (3.3) yield

(3.9) **Corollary.** *Suppose that a state $\mu\in\mathfrak{S}_\theta$ has expected particle density $\mu[X_0=1]=\rho$. Then*

$$s(\mu)-e(\mu,\Phi)\leqq g(\rho,\Phi).$$

Equality holds if and only if $\mu\in\mathfrak{G}_\theta\left(\Phi,-\frac{\partial}{\partial\rho}g(\rho,\Phi)\right)$.

The corollary may be understood as an extension of a standard result on the randomness of product measures. Indeed, if $\Phi \equiv 0$, then $e(\cdot, \Phi) = 0$,

$$g(\rho, \Phi) = -\rho \log \rho - (1-\rho) \log (1-\rho)$$

and $\mathfrak{G}\left(\Phi, -\frac{\partial}{\partial \rho} g(\rho, \Phi)\right)$ consists of the product measure $(1-\rho, \rho)^S$.

4. On the Equivalence of Ensembles

The relation between shift invariant Gibbs states and shift invariant canonical Gibbs states is completely described by

(4.1) **Theorem.** *For any canonical potential* Φ

$$\operatorname{ex} \mathfrak{C}_\theta(\Phi) = \bigcup_{-\infty \leq \varphi \leq +\infty} \operatorname{ex} \mathfrak{G}_\theta(\Phi, \varphi).$$

Proof. Apply (1.10), (2.10), (3.6), and (3.7) and observe that the extreme points of $\mathfrak{C}_\theta$ and $\mathfrak{G}_\theta$ are ergodic and therefore extremal in $\mathfrak{S}_\theta$. □

Because of their homogeneity properties, ergodic states are often called *pure states*, and pure states that describe a particle system with interaction Φ are called (canonical or grand canonical) *pure phases*. According to Dobrushin [2] we say that the parameters Φ and φ induce a (grand canonical, *G*) *phase transition* if $|\operatorname{ex} \mathfrak{G}_\theta(\Phi, \varphi)| > 1$. We see two possibilities of defining a "canonical phase transition". First, let us speak of a *C-I phase transition* with parameters Φ and ρ if the canonical description of a pure state by means of Φ and ρ is not unique, i.e., if the set

$$\mathfrak{C}_\theta(\Phi, \rho) = \{\mu \in \operatorname{ex} \mathfrak{C}_\theta(\Phi) \colon \rho(\cdot) = \rho \; \mu\text{-a.s.}\}$$

contains at least two elements. Examples of *C-I* phase transitions are easily constructed via periodic potentials, see [6].

Theorem (3.7) shows that a *C-I* phase transition at Φ and ρ implies a *G* phase transition at Φ and $-\frac{\partial}{\partial \rho} g(\rho, \Phi)$. The converse is not true since there are models that exhibit a *G* phase transition in the manner that there are exactly two pure phases of distinct particle density, for instance the Ising model [5, 7], some disturbed Ising interactions [1], and some even attractive potentials [16]. These models are characterized by the fact that for some $\rho \in]0, 1[$ there is no pure phase with density ρ. Thus if we want to have a notion of canonical phase transition which is more strongly related to a *G* phase transition we shall have to consider phenomena like

$$|\mathfrak{C}_\theta(\Phi, \rho)| \neq 1$$

which we call a *C-II phase transition* at Φ and ρ. It seems to be hard to find examples for a potential Φ admitting a *G* phase transition at some φ but not a *C-II* phase transition at any ρ. However, even for this weaker notion we have

(4.2) **Proposition.** *A C-II phase transition at* Φ *and* ρ *implies a G phase transition at* Φ *and* $\varphi = -\frac{\partial}{\partial \rho} g(\rho, \Phi)$.

Proof. Assume that $\mathfrak{C}_\theta(\Phi, \rho)=\emptyset$. Then $0<\rho<1$. Denote by ρ_- the smallest and by ρ_+ the largest r such that $\frac{\partial}{\partial\rho} g(r, \Phi)=-\varphi$. Then Theorem (3.7) asserts that the particle density of any $\mu\in\mathfrak{G}_\theta(\Phi, \varphi)$ falls in the interval $[\rho_-, \rho_+]$. Since $\mathrm{ex}\,\mathfrak{G}_\theta(\Phi, \varphi)$ is nonempty, the assumption $\rho_-=\rho_+$ would lead to a contradiction to our hypothesis. Thus $\rho_-<\rho_+$, that is, $g(\cdot, \Phi)$ has a linear segment in its graph. This fact which often serves as a definition of a phase transition (of first kind) is known to imply a G phase transition at Φ and φ, see [5] for a proof. □

Concerning the equivalence of the canonical and the grand canonical description, the results above may be formulated as follows. The Legendre transformation (3.3) Theorem (4.1) is based on asserts that the thermodynamic functions of Φ and φ may be expressed in terms of the thermodynamic functions of Φ and ρ, and vice versa, at least in regions where $\frac{\partial}{\partial\rho} g(\cdot, \Phi)$ is strictly decreasing. For attractive pair potentials these regions coincide with the regions where no grand canonical phase transition occurs [5, 7, 14, 15]. Correspondingly, our theorem states that a pure phase may be described equivalently canonically by Γ_Φ and ρ or grand canonically by $\Lambda_{(\Phi,\varphi)}$ and that both descriptions are unique everywhere that no G phase transition occurs. But we have seen from examples that sometimes a pure phase may be uniquely identified canonically but not grand canonically. On the other hand, it seems to be a very exceptional case that a G phase transition cannot be recognized in the canonical description, and (4.2) excludes the possibility of a false alarm. So in general the canonical description contains at least as much information on the pure phases of a particle system as the grand canonical.

Consider now the class $\mathscr{C}$ of all canonical potentials Φ that do not admit a *C-I* phase transition for any density ρ. $\mathscr{C}$ encloses all canonical potentials that do not exhibit a G phase transition as, for instance, all Φ such that

$$\sum_{0\in A\in\mathscr{S}} (|A|-1)\,|\Phi(A)|<1,$$

and, in the one-dimensional case, all Φ such that

$$\sum_{0\in A\in\mathscr{S}} (\operatorname{diam} A)\,|\Phi(A)|<\infty$$

[2, 7]. Furthermore, $\mathscr{C}$ contains some potentials which allow a G phase transition as noted above. If $\Phi\in\mathscr{C}$, then there is a one-to-one correspondence $\rho\leftrightarrow\nu^\rho$ between a subset of the interval [0, 1] and the extreme points of $\mathfrak{C}_\theta(\Phi)$. Thus the integral representation of a state $\mu\in\mathfrak{C}_\theta(\Phi)$ by a probability measure on $\mathrm{ex}\,\mathfrak{C}_\theta(\Phi)$ may be transformed into a representation via the particle density. More precisely:

(4.3) *Remark. Let $\Phi\in\mathscr{C}$. Then any $\mu\in\mathfrak{C}_\theta(\Phi)$ has a representation*

$$\mu=\int_{[0,1]} \nu^\rho\, P^\mu(d\rho)$$

where P^μ is the distribution of $\rho(\cdot)$ under μ.

Proof. The remark can easily be deduced from the general integral representation mentioned in Section 2, but let us sketch a direct argument. Standard constructions concerning the conditional probabilities $\mu[A|\rho(\cdot)=\rho]$ ($A\in\mathscr{F}$, $0\leqq\rho\leqq 1$)

lead to the existence of shift invariant states $\tilde{\nu}^\rho$ concentrated on the set $\{\rho(\cdot)=\rho\}$ such that

$$\mu=\int \tilde{\nu}^\rho P^\mu(d\rho).$$

By (3.8), $\tilde{\nu}^\rho \in \mathfrak{C}_\theta(\Phi)$ for P^μ-a.a. ρ. Now apply (3.7) to see that for these ρ $\tilde{\nu}^\rho=\nu^\rho$. □

Since the existence of a particle density is basic for the comparison of canonical and grand canonical Gibbs states it is not clear whether in general the sets ex $\mathfrak{C}(\Phi)$ and

$$\bigcup_{-\infty \leqq \varphi \leqq +\infty} \text{ex } \mathfrak{G}(\Phi, \varphi)$$

are comparable by inclusion. In particular, it is an open problem to decide whether the equality $\mathfrak{C}(\Phi)=\mathfrak{C}_\theta(\Phi)$ is true if and only if for all φ $\mathfrak{G}(\Phi,\varphi)=\mathfrak{G}_\theta(\Phi,\varphi)$. In the next section, we give a positive answer to this question in the case of a one-dimensional nearest neighbour potential Φ.

5. Application to Markov Chains and de Finetti's Theorem

In this section we investigate canonical Gibbs states with respect to a one-dimensional nearest neighbour interaction. We show that these are necessarily shift invariant and obtain as an application of the preceding results the characterization of their extreme points as certain Markov chains. This is an extension of de Finetti's theorem (see (5.6) below) to probability measures with a more subtle symmetry property.

Let $S=\mathbb{Z}$. Fix some $a>0$. We consider the canonical nearest neighbour potential Φ_a defined as

$$\Phi_a(A)=\begin{cases}-\log a & \text{if diam } A=1\\ 0 & \text{otherwise}\end{cases} \qquad (A\in\mathscr{S}).$$

The specification Γ_{Φ_a} is given by

$$\gamma_{I,N(\omega)}(\omega|\eta)=Z_{I,N(\omega)}(\eta)^{-1}\, a^{L(\omega)+\eta_{m-1}\,\omega_m+\omega_n\,\eta_{n+1}} \qquad (\omega\in\Omega_I,\ \eta\in\Omega)$$

where $I=[n,m]\in\mathscr{S}$ and

$$L(\omega)=\sum_{x=n}^{m-1}\omega_x\,\omega_{x+1}$$

denotes the number of pairs of particles in ω.

First we characterize $\mathfrak{C}(\Phi_a)$ by a symmetry property.

(5.1) **Proposition.** *Let $\mu\in\mathfrak{S}$. Then $\mu\in\mathfrak{C}(\Phi_a)$ if and only if for all intervals $I\in\mathscr{S}$ the quantity*

$$\mu[X_I=\omega]\, a^{-L(\omega)}$$

depends only on $N(\omega)$ and the value of ω at the endpoints of I.

Proof. The "only if" part results immediately from the definition. Conversely, let I be an interval in S and J any interval that contains I and its neighbours. Then there is a function $z_N(\cdot)$ such that for all $\omega\in\Omega_I$, $\eta\in\Omega$

$$\mu[X_J=\omega\,\eta_{J\setminus I}]\, a^{-L(\omega\eta_{J\setminus I})}=z_{N(\omega)}(\eta_{J\setminus I}).$$

If $z_{N(\omega)}(\eta_{J\setminus I})>0$, this implies that

$$\mu[X_I=\omega|X_{J\setminus I}=\eta_{J\setminus I}, N(X_I)=N(\omega)]=\gamma_{I,N(\omega)}(\omega|\eta)$$

and, in the limit $J\uparrow S$, that $\gamma_I(\omega|\cdot)$ is a version of $\mu[X_I=\omega|\mathscr{G}_I]$. Now let $I\uparrow S$ in order to see that $\mu\in\mathfrak{C}_\infty(\Phi_a)=\mathfrak{C}(\Phi_a)$. □

We use the established symmetry property for a proof that $\mathfrak{C}(\Phi_a)=\mathfrak{C}_\theta(\Phi_a)$. This is done on the base of two lemmas the first of which is probably known.

(5.2) **Lemma.** *Denote by $F^r_{N,L}$ the number of configurations ω on an interval of r consecutive sites such that $N(\omega)=N$ and $L(\omega)=L$. Then $F^r_{0,0}=1$ and, if $N\geqq 1$,*

$$F^r_{N,L}=\binom{N-1}{L}\binom{r-N+1}{N-L}.$$

Since we have no reference we give a proof.

Proof. It is well-known and easily checked that $F^r_{N,0}$ as the N particle part of the r-th Fibonacci number is given by $\binom{r-N+1}{N}$. Obviously, if $N\geqq 1$ then $F^r_{N,L}=0$ unless $L\geqq N-1$ and $r+1\geqq 2N-L$. Now we claim that if $1\leqq L\leqq N-1\leqq r-1$ then

$$LF^r_{N,L}=(N-1)\,F^{r-1}_{N-1,L-1}.$$

Indeed, from any configuration counted by $F^{r-1}_{N-1,L-1}$ we get a configuration ω of $F^r_{N,L}$ as follows: Pick one of its $(N-1)$ particles and split it and its site into a pair of particles on two adjacent sites. Since each of the L pairs in ω may be the result of this doubling process, this transformation is L to 1. Now by induction we obtain

$$F^r_{N,L}=\frac{N-1}{L}\,\frac{N-2}{L-1}\cdots\frac{N-L}{1}\,F^{r-L}_{N-L,0}.\quad \square$$

(5.3) **Lemma.** *Suppose that $\mu\in\mathfrak{S}$ has the property that for all intervals $I\in\mathscr{S}$ the marginal distributions $\mu[X_I=\omega]$ depend only on $N(\omega)$, $L(\omega)$, and the values of ω at the ends of I. Fix some $n\geqq 1$ and a configuration $\zeta\in\Omega_{[0,n-1]}$ with $\zeta_0=\zeta_{n-1}=1$. Then for any $k\geqq 0$ there are configurations ω^j and $\tilde\omega^j$ $(0\leqq j\leqq k)$ on the interval $I(n,k)=[-k,n+k]$ such that $N(\omega^j)=N(\tilde\omega^j)=N(\zeta)+k$, $L(\omega^j)=L(\tilde\omega^j)=L(\zeta)+k-j$, $\omega^j_{-k}+\omega^j_{n+k}=1$, $\tilde\omega^j_{-k}+\tilde\omega^j_{n+k}=1$, and*

$$\begin{aligned}\mu[X_{[1,n]}=\theta_1\zeta]-\mu[X_{[0,n-1]}=\zeta]\\ =\sum_{j=0}^{k}\binom{k}{j}\{\mu[X_{I(n,k)}=\omega^j]-\mu[X_{I(n,k)}=\tilde\omega^j]\}.\end{aligned}$$

Proof. For the induction proof we need as an additional property that $\omega^j_{-k}=\tilde\omega^j_{n+k}=0$, $\omega^j_{n+k}=\tilde\omega^j_{-k}=1$ if j is even and the same with 0 and 1 interchanged if j is odd. We start with the equation

$$\begin{aligned}\mu[X_{[1,n]}=\theta_1\zeta]-\mu[X_{[0,n-1]}=\zeta]\\ =\mu[X_{[0,n]}=0(\theta_1\zeta)]+\mu[X_{[0,n]}=1(\theta_1\zeta)]\\ -\mu[X_{[0,n]}=\zeta 0]-\mu[X_{[0,n]}=\zeta 1].\end{aligned}$$

The second and the fourth term cancel owing to our assumptions. This settles the case $k=0$. Now assume that for some $k \geqq 0$ the refined statement is true. For any $\omega \in \Omega_{I(n,k)}$ we write

$$\mu[X_{I(n,k)}=\omega]=\mu[X_{I(n,k+1)}=0\,\omega\,0]+\mu[X_{I(n,k+1)}=1\,\omega\,1] \\ +\mu[X_{I(n,k+1)}=0\,\omega\,1]+\mu[X_{I(n,k+1)}=1\,\omega\,0]$$

and deduce from the symmetry hypothesis that

$$\begin{aligned}\mu[X_{I(n,k)}=\omega^j]-\mu[X_{I(n,k)}=\tilde{\omega}^j] &=\mu[X_{I(n,k+1)}=1\,\omega^j 0]-\mu[X_{I(n,k+1)})=0\,\tilde{\omega}^j 1] \\ &\quad+\mu[X_{I(n,k+1)}=0\,\omega^j 1]-\mu[X_{I(n,k+1)}=1\,\tilde{\omega}^j 0]\end{aligned}$$

and for even $j<k$

$$\mu[X_{I(n,k+1)}=1\,\omega^j 0]=\mu[X_{I(n,k+1)}=1\,\omega^{j+1}\,0]$$

and similar identities if j is odd or ω^j is replaced by $\tilde{\omega}^j$. Setting

$$\alpha^j=\begin{cases}0\;\omega^j\;1 & \text{if } j \quad \text{even } \leqq k\\ 1\;\omega^j\;0 & \text{if } j \quad \text{odd } \leqq k\\ 0\;\omega^k\;1 & \text{if } j=k+1 \text{ even}\\ 1\;\omega^k\;0 & \text{if } j=k+1 \text{ odd}\end{cases}$$

and similarly $\tilde{\alpha}^j$ with 0 and 1 exchanged we see that α^j and $\tilde{\alpha}^j$ $(0 \leqq j \leqq k+1)$ are of the required form, and with the abbreviation

$$\mu_j=\mu[X_{I(n,k+1)}=\alpha^j]-\mu[X_{I(n,k+1)}=\tilde{\alpha}^j]$$

we have

$$\mu[X_{[1,n]}=\theta_1\,\zeta]-\mu[X_{[0,n-1]}=\zeta]=\sum_{j=0}^{k}\binom{k}{j}[\mu_j+\mu_{j+1}]=\sum_{j=0}^{k+1}\binom{k+1}{j}\mu_j. \quad \square$$

(5.4) **Theorem.** (a) *$\mathfrak{C}(\Phi_a)$ consists only of shift invariant and thereby reversible states.*

(b) *For any $\rho \in [0,1]$ there is one and only one $\nu^\rho \in \operatorname{ex}\mathfrak{C}(\Phi_a)$ with a.s. constant particle density ρ. ν^ρ is the ergodic Markov measure for the transition matrix*

$$M=\begin{pmatrix}1-\rho/b & \rho/b\\ (1-\rho)/b & 1-(1-\rho)/b\end{pmatrix}$$

where $b=\frac{1}{2}+[\frac{1}{4}+\rho(1-\rho)(a-1)]^{1/2}$.

(c) *Any $\mu \in \mathfrak{C}(\Phi_a)$ has a representation*

$$\mu=\int_{[0,1]}\nu^\rho\,P^\mu(d\rho),$$

where P^μ is the distribution of $\rho(\cdot)$ under μ.

Proof. (a) Let $\mu\in\mathfrak{C}(\Phi_a)$. Then (5.1) and (5.3) show that for all $n, k\geqq 1$ and all $\zeta\in\Omega_{[0,n-1]}$ with $\zeta_0=\zeta_{n-1}=1$ and $N(\zeta)=N$, $L(\zeta)=L$

$$|\mu[X_{[1,n]}=\theta_1\zeta]-\mu[X_{[0,n-1]}=\zeta]|$$
$$\leqq\sum_{j=0}^{k}\binom{k}{j}\int_{\Omega_{I(n,k),N+k}}d\mu\, 2\max_{\omega:\,L(\omega)=L+k-j}\gamma_{I(n,k),N+k}(\omega|\cdot).$$

Setting $a_*=\max(a, 1/a)$ we see that the maximum under the integral is bounded by

$$a_*^2\, a^{L+k-j}/a_*^{-2}\sum_{\omega\in\Omega_{I(n,k),N+k}}a^{L(\omega)}\leqq a_*^4/F^{n+2k+1}_{N+k,L+k-j}.$$

Now apply (5.2) and observe that since $0\leqq N-L-1\leqq N-1$ and $N-L-1\leqq n-N$ the right-hand side may be estimated by

$$a_*^4\Big/\binom{k}{j}\binom{k+2}{j+1}.$$

Thus the difference under consideration has the upper bound

$$2a_*^4\sum_{j=0}^{k}\binom{k+2}{j+1}^{-1}$$

vanishing in the limit $k\to\infty$.

Summing up over ζ we see that for all $A\in\mathscr{S}$ and $x\in S$

$$\mu[X_A=\mathbb{1}]=\mu[X_{A+x}=\mathbb{1}].$$

Now the shift invariance of μ follows from the inclusion – exclusion formula

$$\mu[X_V=\omega]=\sum_{V(\omega)\subset A\subset V}(-1)^{|A\setminus V(\omega)|}\mu[X_A=\mathbb{1}]\qquad(\omega\in\Omega_V,\ V\in\mathscr{S})$$

where $V(\omega)=\{x\colon \omega_x=1\}$. If V is an interval, a reversing of the order of coordinates of ω entails on the right-hand side only shifts of the sets A proving the reversibility of μ.

(b) An application of part (a) and Theorem (3.7) yields that for any $\mu\in\operatorname{ex}\mathfrak{C}(\Phi_a)$ with particle density ρ there is a φ uniquely determined by ρ such that

$$\mu\in\operatorname{ex}\mathfrak{G}_\theta(\Phi_a, \varphi).$$

Since for all φ $|\mathfrak{G}(\Phi_a, \varphi)|=1$ (see [2, 7]) it is enough to show that the Markov measure described in the statement is well defined, has density ρ, and is a Gibbs state for Φ_a and some φ. It is easily checked that M is a stochastic matrix and admits $(1-\rho, \rho)$ as an invariant probability vector. To prove that it is a Gibbs state it suffices to realize that μ has the right one-point conditional probabilities $\lambda_{\{x\}}(\cdot|\cdot)$ for an appropriate φ [2, 7]. This is done by an elementary calculation of the ratios

$$M(i, 1)\, M(1, j)/M(i, 0)\, M(0, j)\qquad(i, j\in\{0, 1\})$$

which we omit here. The proper φ is given by

$$\varphi=-\log[(b-\rho)^2/\rho(1-\rho)]. \tag{5.5}$$

Part (c) is an application of (4.3). □

Theorem (5.4) is an extension of de Finetti's theorem concerning exchangeable 0-1 variables to Markov chains. Indeed, if $a=1$ then by (5.1) $\mathfrak{C}(\Phi_1)$ consists exactly of all symmetric states. Furthermore, if $a=1$ then $b=1$ so that ν^ρ is the product measure $(1-\rho, \rho)^S$. So in this case (5.4) reads as follows:

(5.6) **Corollary** (de Finetti). *Any probability measure μ on $\{0,1\}^{\mathbb{Z}}$ which is invariant under permutations of finitely many coordinates can be obtained from the product measures $(1-\rho, \rho)^{\mathbb{Z}}$ $(0 \leqq \rho \leqq 1)$ by a randomization of the parameter ρ.*

6. "Microcanonical" Markov States

In this section we proceed in investigating one-dimensional nearest neighbour systems, but now we go one step further and treat, in addition to the particle density ρ, the nearest neighbour coupling constant a as a free parameter. The resulting set $\mathfrak{M}$ of states is determined only by the property that only nearest neighbours interdepend.

If the particle density and the pair correlation of a state μ in $\mathfrak{M}$ are a.s. constant then it is possible to associate with μ a fixed coupling constant a, and a determines already the interaction energy of μ. For this reason a description of a state by ρ, a, and the property of belonging to $\mathfrak{M}$ is similar to that in the microcanonical thermodynamic ensemble.

Proposition (5.1) asserts that for any $\mu \in \bigcup_{a>0} \mathfrak{C}(\Phi_a)$ the marginal distribution $\mu[X_I=\omega]$ in an interval I depends only on $N(\omega)$, $L(\omega)$ and the values of ω at the ends of I. By the reversibility of μ, in fact only the sum of the values of ω at the ends of I is essential. Let us give a more intuitive description of this symmetry property. We say that a configuration ω on a finite interval I is of the *type* (N, L, l) if its particle number, its number of pairs of particles and of pairs of blanks are given by N, L, and l, respectively, i.e., if $N(\omega)=N$, $L(\omega)=L$, and $L(\mathbb{1}-\omega)=l$.

(6.1) *Remark.* Two configurations ω, ζ on an interval $[m,n] \in \mathcal{S}$ are of the same type if and only if $N(\omega)=N(\zeta)$, $L(\omega)=L(\zeta)$, and $\omega_m+\omega_n=\zeta_m+\zeta_n$.

Proof. This follows immediately from the equation

$$2N(\omega) = -L(\mathbb{1}-\omega) + L(\omega) + (n-m) + \omega_m + \omega_n$$

which is a consequence of the definition of $L(\cdot)$. □

In this section we investigate the set $\mathfrak{M}$ of all states $\mu \in \mathfrak{S}$ whose marginal distributions $\mu[X_I=\omega]$ in all finite intervals I depend only on the type of ω. We call any $\mu \in \mathfrak{M}$ a *microcanonical Markov state.*

(6.2) *Remark.* (a) $\mathfrak{M}$ is convex and weakly compact.

(b) $\mathfrak{M} \subset \mathfrak{S}_\theta$.

(c) Any shift invariant Markovian probability measure is an extreme point of $\mathfrak{M}$.

Proof. (a) is evident. To prove (b), apply (6.1) and (5.3) and observe that the configurations ω^j, $\tilde{\omega}^j$ in (5.3) are of the same type. This implies the shift invariance of any $\mu \in \mathfrak{M}$ just as in the proof of (5.4). Finally, it is easily checked that every shift invariant Markov chain belongs to $\mathfrak{M}$. They are ergodic, that is, extremal in the larger set $\mathfrak{S}_\theta$. □

Our aim is to prove the converse of (c): that the extreme points of $\mathfrak{M}$ are shift invariant Markov measures. The methods are similar to those in Section 3. Fix again a sequence V of intervals in $\mathscr{S}$ whose length tends to infinity.

Now let us introduce the 0-1 pair correlation function

$$\sigma(\omega)=\limsup_{V\uparrow S}|V|^{-1}\sum_{x\in V}\omega_x(1-\omega_{x+1})$$
$$=\limsup_{V\uparrow S}|V|^{-1}[N(\omega_V)-L(\omega_V)] \quad (\omega\in\Omega).$$

This is a more convenient parameter than the 1-1 correlation function $\rho(\omega)-\sigma(\omega)$. If $\mu\in\mathfrak{S}_\theta$, then for μ-a.a. $\omega\in\Omega$ $\sigma(\omega)$ is in fact a limit and coincides with $E_\mu[X_0(1-X_1)|\mathscr{I}](\omega)$. The role of the Helmholtz free energy is played by the function

$$m(\rho,\sigma)=(\rho-\sigma)\log\frac{\rho}{\rho-\sigma}+\sigma\log\frac{\rho(1-\rho)}{\sigma^2}$$
$$+(1-\rho-\sigma)\log\frac{1-\rho}{1-\rho-\sigma} \quad (0\leqq\sigma\leqq\rho;\ \sigma\leqq 1-\rho).$$

(6.3) **Lemma.** *$m(\cdot,\cdot)$ is a strictly concave function.*

Proof. According to [8] this can be proved by calculating the second derivatives of $m(\cdot,\cdot)$ and observing that they form a negative definite matrix since

$$\frac{\partial^2}{\partial\sigma^2}m(\rho,\sigma)=-\frac{1}{\rho-\sigma}-\frac{2}{\sigma}-\frac{1}{1-\rho-\sigma}<0,$$

$$\frac{\partial^2}{\partial\sigma^2}m(\rho,\sigma)\frac{\partial^2}{\partial\rho^2}m(\rho,\sigma)-\left[\frac{\partial^2}{\partial\sigma\,\partial\rho}m(\rho,\sigma)\right]^2=\frac{2\rho(1-\rho)+1}{\rho(1-\rho)(\rho-\sigma)(1-\rho-\sigma)}>0.$$

(6.4) **Lemma.** *$m(\rho,\sigma)$ is the limit of $r^{-1}\log F^r_{N,L}$ if $r\to\infty$ and $N/r\to\rho$, $(N-L)/r\to\sigma$ such that eventually $F^r_{N,L}\geqq 1$. In particular, $m(\cdot,\cdot)\geqq 0$.*

Proof. This is a straightforward application of (5.2) and Stirling's formula. □

The variational characterization of $\mathfrak{M}$ is based on the affine functional

$$m(\mu)=E_\mu[m(\rho(\cdot),\sigma(\cdot))]-s(\mu)$$

on $\mathfrak{S}_\theta$.

(6.5) **Proposition.** *For any $\mu\in\mathfrak{S}_\theta$, $m(\mu)\geqq 0$. If $\mu\in\mathfrak{M}$ then $m(\mu)=0$.*

Proof. For any interval $V=[m,n]\in\mathscr{S}$ and $i,j\in\{0,1\}$ consider the set

$$\Omega_{V,N,L}(ij)=\{\omega\in\Omega_V\colon N(\omega)=N,\ L(\omega)=L,\ \omega_m=i,\ \omega_n=j\}.$$

Its cardinality $Z_{V,N,L}(ij)$ plays the role of a microcanonical partition function. Obviously, we have

$$F^{|V|-4}_{N-i-j,L}\leqq Z_{V,N,L}(ij)\leqq F^{|V|}_{N,L} \quad (N\geqq i+j).$$

Thus we conclude from (6.4) that if $N/|V| \to \rho$, $(N-L)/|V| \to \sigma$, then

$$|V|^{-1} \log Z_{V,N,L}(ij)$$

tends to $m(\rho, \sigma)$ independently of i and j.

For any $\mu \in \mathfrak{S}_\theta$ consider now the expression

$$m_V(\mu) = \int \mu(d\eta) \log \frac{\mu[X_V = \eta_V]\, Z_{V,N(\eta_V),L(\eta_V)}(\eta_m\, \eta_n)}{\mu[X_V \in \Omega_{V,N(\eta_V),L(\eta_V)}(\eta_m\, \eta_n)]}.$$

If $\mu \in \mathfrak{M}$, then the ratio under the logarithm a.s. equals 1 so that $m_V(\mu) = 0$. For arbitrary μ we may write $m_V(\mu)$ as

$$\sum_{i,j=0}^{1} \sum_{N=1}^{|V|} \sum_{L=0}^{N-1} \mu[X_V \in \Omega_{V,N,L}(ij)] \sum_{\omega \in \Omega_{V,N,L}(ij)} \mu[X_V = \omega \,|\, \Omega_{V,N,L}(ij)]$$
$$\cdot \log \frac{\mu[X_V = \omega \,|\, \Omega_{V,N,L}(ij)]}{Z_{V,N,L}(ij)^{-1}}.$$

The inner sum has the form of a relative entropy and is therefore nonnegative. Hence it is enough to show that $|V|^{-1} m_V(\mu)$ converges to $m(\mu)$. To this end write the logarithm of the product as the sum of the logarithms. Then the first term tends to $-s(\mu)$. Since $|V|^{-1} \log Z_{V,N,L}(\cdot)$ is bounded uniformly in V, N, and L, the ergodic theorem guaranties that

$$\int \mu(d\eta) |V|^{-1} \log Z_{V,N(\eta_V),L(\eta_V)}(\eta_m\, \eta_n)$$

tends to $\int \mu(d\eta)\, m(\rho(\eta), \sigma(\eta))$. The remaining term can be written in the form of an entropy

$$-|V|^{-1} \sum_{i,j,N,L} \mu[X_V \in \Omega_{V,N,L}(ij)] \log \mu[X_V \in \Omega_{V,N,L}(ij)]$$

and has therefore the bound $|V|^{-1} \log[2|V|(|V|+1)]$ vanishing in the limit $|V| \to \infty$. □

Now we study states $\mu \in \mathfrak{S}_\theta$ with the properties $m(\mu) = 0$, $\rho(\cdot) = \rho$, $\sigma(\cdot) = \sigma$ μ-a.s. It is natural to guess that such a state is the unique Markovian probability measure with density ρ and 0-1 correlation σ having the stochastic matrix

$$M^{\rho,\sigma} = \begin{cases} \begin{pmatrix} 1-\sigma/(1-\rho) & \sigma/(1-\rho) \\ \sigma/\rho & 1-\sigma/\rho \end{pmatrix} & \text{if } 0 < \rho < 1 \\ \begin{pmatrix} 1-\rho & \rho \\ 1-\rho & \rho \end{pmatrix} & \text{if } \rho \in \{0, 1\} \end{cases} \quad (\sigma \leqq \rho \leqq 1-\sigma)$$

as transition matrix. We denote this Markov measure by $\nu^{\rho,\sigma}$. Obviously, $\nu^{\rho,\rho(1-\rho)} = (1-\rho, \rho)^S$.

(6.6) **Proposition.** *Suppose that for some* $\mu \in \mathfrak{S}_\theta$ $m(\mu) = 0$ *and* $\rho(\cdot) = \rho$, $\sigma(\cdot) = \sigma$ μ-*a.s. Then* $\mu = \nu^{\rho,\sigma}$. *In particular,* μ *is ergodic.*

Proof. Consider first the main case $0 < \sigma < \rho$, $\sigma < 1-\rho$. Then we may find some $a > 0$ and $\varphi \in \mathbb{R}$ such that $\{\nu^{\rho,\sigma}\} = \mathfrak{G}(\Phi_a, \varphi)$. Indeed, setting

$$a = 1 + (\rho - \sigma - \rho^2)\, \sigma^{-2} > 0$$

and defining φ from a and ρ via (5.5), namely

$$\varphi = -\log\left[\frac{\rho}{1-\rho}\left(\frac{1-\rho-\sigma}{\sigma}\right)^2\right]$$

we find that $M^{\rho,\sigma}$ coincides with the matrix M considered in (5.4). Now observe that $P(\Phi_a, \varphi) = -f(\nu^{\rho,\sigma}; \Phi_a, \varphi) = -\log[1-\sigma/(1-\rho)]$ (see [7]). Furthermore, with this definition of φ and a we have

$$0 = m(\rho,\sigma) - s(\mu) = -\varphi\rho - (\rho-\sigma)\log a + P(\Phi_a, \varphi) - s(\mu) = f(\mu; \Phi_a, \varphi) + P(\Phi_a, \varphi)$$

proving that $\mu \in \mathfrak{G}(\Phi_a, \varphi)$.

Consider now the exceptional cases. If $\rho \in \{0, 1\}$, then $\sigma = 0$ and $\mu = \varepsilon_\rho = \nu^{\rho,\sigma}$. If $\sigma = 0$ then for all $m \leqq n$, $\rho = \mu[X_{[m,n]} = \mathbb{1}] = 1 - \mu[X_{[m,n]} = 0]$ so that $\mu = \rho\varepsilon_1 + (1-\rho)\varepsilon_0$. Since $\rho(\cdot) = \text{const}$ μ-a.s. we must have $\rho = 0$ or 1. Some more work will be necessary for the case $0 < \sigma = \rho < 1$ which is non-trivial, but has infinite specific energy since then $a = 0$. So we have to argue more carefully. First observe that μ is carried by

$$\Omega_0 = \{\eta \in \Omega \colon L(\eta_{[-n,n]}) = 0 \text{ for all } n \geqq 0\}.$$

For $n \geqq 0$ consider now

$$s_n(\mu \mid \nu^{\rho,\rho}) = \frac{1}{n+1} \int \mu(d\eta) \log \frac{\mu[X_{[0,n]} = \eta_{[0,n]}]}{\nu^{\rho,\rho}[X_{[0,n]} = \eta_{[0,n]}]}.$$

If $\eta \in \Omega_0$ then

$$\nu^{\rho,\rho}[X_{[0,n]} = \eta_{[0,n]}] = \nu^{\rho,\rho}[X_0 = \eta_0]\left(\frac{\rho}{1-\rho}\right)^{N(\eta_{[1,n]})}\left(\frac{1-2\rho}{1-\rho}\right)^{L(1-\eta_{[0,n]})} > 0.$$

The formula in the proof of (6.1) therefore shows that for μ-a.a. η

$$\frac{1}{n+1}\log \nu^{\rho,\rho}[X_{[0,n]} = \eta_{[0,n]}]$$

converges (boundedly) to $-m(\rho, \rho)$ if $n \to \infty$. Thus

$$s_n(\mu \mid \nu^{\rho,\rho}) \to -s(\mu) + m(\rho,\rho) = 0.$$

But writing $s_n(\mu \mid \nu^{\rho,\rho})$ in the form

$$\frac{1}{n+1}\int \mu(d\eta) \sum_{j=0}^{n} \log \frac{\mu[X_j = \eta_j \mid \mathscr{F}_{[0,j-1]}](\eta)}{\nu^{\rho,\rho}[X_j = \eta_j \mid \mathscr{F}_{[0,j-1]}](\eta)}$$

we see just as in [3, Prop. 3.2] that its vanishing limit is given by

$$\int \mu(d\eta) \sum_{\omega_0 \in \Omega_{\{0\}}} \mu[X_0 = \omega_0 \mid \mathscr{F}_{]-\infty,0[}](\eta) \log \frac{\mu[X_0 = \omega_0 \mid \mathscr{F}_{]-\infty,0[}](\eta)}{M^{\rho,\rho}(\eta_{-1}, \omega_0)}.$$

From this we get that μ is Markovian with transition matrix $M^{\rho,\rho}$ and thereby that $\mu = \nu^{\rho,\rho}$.

The final case $0 < \sigma = 1 - \rho < 1$ can be reduced to the preceding one by the "spin flip" $\omega \to \mathbb{1} - \omega$. □

(6.7) **Theorem.** (a) *A state* $\mu \in \mathfrak{S}_\theta$ *is a microcanonical Markov state if and only if* $m(\mu)=0$.

(b) *Any* $\mu \in \mathfrak{M}$ *is uniquely determined by the common distribution* P^μ *of* $(\rho(\cdot), \sigma(\cdot))$ *under* μ. *More precisely,* μ *has a representation*

$$\mu = \int_{[0,1]^2} \nu^{\rho,\sigma} P^\mu(d\rho, d\sigma).$$

In particular, $\mathrm{ex}\,\mathfrak{M} = \{\nu^{\rho,\sigma}: 0 \leqq \sigma \leqq \rho; \sigma \leqq 1-\rho\}$.

(c) *Suppose that for some* $\mu \in \mathfrak{S}_\theta$ $E_\mu[\rho(\cdot)]=\rho$ *and* $E_\mu[\sigma(\cdot)]=\sigma$. *Then* $s(\mu) \leqq m(\rho, \sigma)$. *The maximum value* $m(\rho, \sigma)$ *of the entropy is attained exactly by the Markov measure* $\nu^{\rho,\sigma}$.

Proof. Fix some $\mu \in \mathfrak{S}_\theta$ with $m(\mu)=0$. Then the conditional probability kernel $\mu[A \mid \rho(\cdot)=\rho, \sigma(\cdot)=\sigma]$ yields for P^μ-a.a. (ρ, σ) states $\tilde{\nu}^{\rho,\sigma} \in \mathfrak{S}_\theta$ with $\tilde{\nu}^{\rho,\sigma}[\rho(\cdot)=\rho, \sigma(\cdot)=\sigma]=1$ and $\mu = \int \tilde{\nu}^{\rho,\sigma} P^\mu(d\rho, d\sigma)$. Since

$$0 = m(\mu) = \int m(\tilde{\nu}^{\rho,\sigma}) P^\mu(d\rho, d\sigma)$$

we have $m(\tilde{\nu}^{\rho,\sigma})=0$ and thereby $\tilde{\nu}^{\rho,\sigma}=\nu^{\rho,\sigma}$ P^μ-a.s. Combined with (6.5) this proves (a) and (b). The first part of (c) follows from (a) and (6.3) by Jensen's inequality. If $0 \leqq m(\mu) \leqq m(\rho, \sigma) - s(\mu) = 0$, then again Jensen's inequality asserts that $\rho(\cdot)=\rho$, $\sigma(\cdot)=\sigma$ μ-a.s. Now apply (6.6). □

Acknowledgement. I am grateful to H. Rost for stimulating my interest in canonical specifications and for a discussion of the results.

References

1. Capocaccia, D., Cassandro, M., Ciccotti, G.: Equilibrium states of an Ising ferromagnet in the low temperature region. Comm. Math. Phys. **29**, 31–42 (1973)
2. Dobrushin, R.L.: Description of a random field by means of its conditional probabilities and conditions of its regularity. Theo. Probability Appl. **13**, 197–224 (1968)
   Dobrushin, R.L.: Gibbsian random fields for lattice systems with pair interactions. Functional Anal. Appl. **2**, 281–301 (1968)
   Dobrushin, R.L.: The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. Functional Anal. Appl. **2**, 302–312 (1968)
3. Föllmer, H.: On entropy and information gain in random fields. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 207–217 (1973)
4. Föllmer, H.: Phase transition and Martin boundary. To appear in: Seminaire de Probabilités IX (Strasbourg)
5. Georgii, H.O.: Phasenübergang 1. Art bei Gittergasmodellen. Lect. Notes Physics **16**. Berlin, Heidelberg, New York: Springer 1972
6. Georgii, H.O.: Two remarks on extremal equilibrium states. Comm. Math. Phys. **32**, 107–118 (1973)
7. Georgii, H.O.: Stochastische Felder und ihre Anwendung auf Interaktionssysteme. Lecture Notes, Institut für Angewandte Mathematik der Universität Heidelberg 1974
8. Hardy, G.H., Littlewood, J.E., Polya, G.: Inequalities. Cambridge: Univ. Press 1934
9. Hewitt, E., Savage, L.J.: Symmetric measures on Cartesian products. Trans. Amer. Math. Soc. **80**, 470–501 (1955)
10. Holley, R.: A class of interactions in an infinite particle system. Advances in Math. **5**, 291–309 (1970)
    Holley, R.: Pressure and Helmholtz free energy in a dynamic model of a lattice gas. Proc. Sixth Berkeley Symp. Prob. math. Statist. **3**, 565–578 (1971)

11. Holley, R.: Free energy in a Markovian model of a lattice spin system. Comm. Math. Phys. **23**, 87–99 (1971)
Holley, R.: An ergodic theorem for interacting systems with attractive interaction. Z. Wahrscheinlichkeitstheorie verw. Gebiete **24**, 325–334 (1972)
12. Lanford, O.E., Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. Comm. Math. Phys. **13**, 194–215 (1969)
13. Liggett, Th.M.: A characterization of the invariant measures for an infinite particle system with interactions II. Trans. Amer. Math. Soc. **198**, 201–213 (1974)
14. Preston, C.: Gibbs states on countable sets. Cambridge Tracts in Mathematics **68**. London: Cambridge Univ. Press 1974
15. Preston, C.: Notes on random fields (to be published)
16. Slawny, J.: Analyticity and uniqueness for spin 1/2 classical ferromagnetic lattice systems at low temperatures. Comm. Math. Phys. **34**, 271–296 (1973)
17. Spitzer, F.: Interaction of Markov Processes. Advances in Math. **5**, 246–290 (1970)
18. Spitzer, F.: Introduction aux Processus de Markov a Parametres dans Z_ν. Ecole d'Eté de Probabilités de Saint-Flour III-1973, Lect. Notes Math. **390**. Berlin, Heidelberg, New York: Springer 1974

Hans-Otto Georgii
Institut für Angewandte Mathematik
der Universität
D-6900 Heidelberg
Im Neuenheimer Feld 5
Federal Republic of Germany

(Received March 5, 1975)