Z. Wahrscheinlichkeitstheorie verw. Gebiete
58, 69–85 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Almost Sure Entropy and the Variational Principle for Random Fields with Unbounded State Space

Hansrudolf Künsch

Mathematikdepartement, ETH-Zentrum, CH-8092 Zürich, Switzerland

0. Introduction

A *random field* is a stochastic process $(X_i)_{i \in \mathbb{Z}^d}$ indexed by the integer lattice $\mathbb{Z}^d$. A random field is called *Gibbsian* if the conditional distributions of X_i, $i \in V$, given the values X_j, $j \notin V$, can be written in terms of an energy function for all finite $V \subseteq \mathbb{Z}^d$ (cf. e.g. [8]). If several Gibbsian fields with the same energy exist, it is interpreted as *phase transition.* The Gibbs variational principle says that among all stationary random fields the Gibbsian fields are characterized by the fact that they minimize a suitable *free energy* which is a justification for calling the Gibbsian fields equilibria. It was proved by Lanford and Ruelle [4] if the X_i take only the two values 0 or 1. Föllmer [1] discovered a connection of this with information theoretic quantities: He proved that the *information gain* of a stationary field over a Gibbsian field is the difference of the free energies. The variational principle is then equivalent to the claim that two stationary fields have information gain zero iff they have the same conditional distributions for finite sets given the outside.

If the state space is unbounded, new difficulties arise because in general uniform bounds are no more available. Usually one works here with the *superstability* and *regularity* assumptions (see [11, 12]) which give powerful estimates for the conditional distributions (Ruelle [12]). With the help of these estimates Lebowitz and Presutti [5] proved the existence of the so called pressure, and with this result Pirlot [7] showed that Gibbsian fields minimize the free energy and that two Gibbsian fields with the same energy have information gain zero. A general result of Preston [8] says in our situation that if the information gain of a stationary random field with respect to a Gibbsian field with energy of finite range is zero, it must be Gibbsian with the same energy. The case of infinite range is still open.

This paper contains the following results: We prove almost sure and L_1-convergence of the entropy for Gibbsian fields (d-dimensional version of the theorem of McMillan-Breiman) which gives us also an almost sure version of the variational principle. Secondly we show that the formula information gain equal difference of free energies holds also for unbounded state space. This formula makes it possible to prove Prestons result about information gain mentionned before also in cases of infinite range of the energy. Moreover our approach gives new proofs for the existence of pressure and of Pirlots results. It consists essentially in proving that the conditional density approximates in some sense the absolute density (Sect. 3). In the case of finite state space this is not very hard to prove (Föllmer [1], formula (4.22)), but in our case it requires rather complicated estimates and a condition somewhat stronger than regularity. In Sect. 2 we prove the almost sure and L_1-convergence of the energy with a version of the ergodic theorem for nonadditive functions. Section 4 contains the main results.

The author wishes to express his gratitude to H. Föllmer for suggesting the problem and helping with valuable suggestions. Thanks are also due to Gallavotti for a stimulating discussion.

1. Definitions and Assumptions

We consider the configuration space $X=(\mathbb{R}^n)^{\mathbb{Z}^d}$, i.e. the set of all functions $x\colon \mathbb{Z}^d \to \mathbb{R}^n$. x_i will be the value at the site i. Let $\mathscr{B}$ be the σ-field generated by the projections $x \to x_i$, $i \in \mathbb{Z}^d$. A $\mathbb{R}^n$-valued *random field* is then a probability measure on $(X, \mathscr{B})$. A random field μ will be called *stationary* if μ is invariant under all shifttransformations θ_i, $i \in \mathbb{Z}^d$, where $(\theta_i x)_j = x_{j+i}$. We will call μ *tempered* (in the sense of Ruelle [11]) if

(1.1) $$\mu\{\bigcup_N \bigcap_n (\sum_{|i| \leq n} x_i^2 \leq N^2 (2n+1)^d)\} = 1, \qquad \text{where } |.|$$

is the norm on $\mathbb{Z}^d$ defined by $|i| = \max(|i_1|, \ldots, |i_d|)$.

For any $D \subseteq \mathbb{Z}^d$ $\mathscr{B}(D)$ will be the σ-field generated by the projections $x \to x_i$, $i \in D$. $\mathscr{V}$ denotes the set of all finite subsets of $\mathbb{Z}^d$. The letters V and W will be reserved for elements of $\mathscr{V}$. $|V|$ denotes the cardinality of the set $V \in \mathscr{V}$.

We start with an *energy* U, i.e. a family of measurable functions U_V, $V \in \mathscr{V}$, $U_V\colon (\mathbb{R}^n)^V \to \mathbb{R}$ which satisfies the following consistency and normalizing conditions

(1.2) $$U_V(x) = U_W(x) \quad \text{if } W \subseteq V \text{ and } x_i = 0 \text{ for all } i \in V \setminus W, \qquad U_i(0) = 0.$$

The corresponding *interaction* I is defined by

(1.3) $$I_{V,W}(x) = U_{V \cup W}(x) - U_V(x) - U_W(x) \qquad (V \cap W = \emptyset),$$

and the *potential* Φ by

(1.4) $$\Phi_V(x) = \sum_{W \subset V} (-1)^{|V|-|W|} U_W(x).$$

It follows from (1.2) and (1.4) that

(1.5) $$U_V(x) = \sum_{W \subset V} \Phi_W(x) \quad \text{and} \quad \Phi_V(x) = 0 \text{ if } x_i = 0 \text{ for some } i \in V.$$

Usually we will work rather with the interaction than with the potential.

We will always suppose that the energy is *stationary*, i.e.

(1.6) $$U_V(x)=U_{V+i}(\theta_{-i}x) \qquad (i\in\mathbb{Z}^d,\ V\in\mathscr{V}).$$

As the need arises, we will impose the following conditions on U:

(1.7) *Definition.* An energy U is called *superstable* if there exist $A>0$ and $B\in\mathbb{R}$ such that $U_V(x)\geqq A\sum_{i\in V} x_i^2+B|V|$ for all $V\in\mathscr{V}$. It is called *regular (super-regular)* if there exists $\psi:\mathbb{N}\to\mathbb{R}_+$ monotonically decreasing with $\sum_n \psi(n)\,n^{d-1}<\infty$ such that

$$|I_{V,W}(x)|\leqq\tfrac{1}{2}\sum_{i\in V,\,j\in W}\psi(|i-j|)(x_i^2+x_j^2),$$

respectively

$$|I_{V,W}(x)|\leqq\sum_{i\in V,\,j\in W}\psi(|i-j|)\,|x_i|\,|x_j|.$$

Superstability and regularity were introduced by Ruelle [12]. Superregularity which is slightly stronger than regularity is new here, it will be needed for our main result (Theorem 3.14).

With the same methods as in [8], p. 95–100, it can be shown that to every superstable energy there exists a consistent family of *conditional distributions with respect to* $\mathscr{B}(V^c)$, $V\in\mathscr{V}$, which are of the following form

$$=\pi^V(x\,|\,y)\,d\omega_V(x)\ \ (x\in(\mathbb{R}^n)^V,\ y\in(\mathbb{R}^n)^{V^c}) \qquad \text{if}\ \ y\in R_V$$
$$=0 \qquad \text{if}\ \ y\notin R_V.$$

Here $d\omega_V(x)$ is a product *reference measure* $\prod_{i\in V} d\omega(x_i)$ on $(\mathbb{R}^n)^V$ with

(1.8) $$\int\exp(-cx^2)\,d\omega(x)<\infty \qquad (c>0),$$

R_V is an element of $\mathscr{B}(V^c)$, the set of nice *boundary conditions* y for which $I_{V,V^c}(xy)$ is well defined, and

(1.9) $$\pi^V(x\,|\,y)=Z_V(y)^{-1}\exp(-U_V(x)-I_{V,V^c}(xy))$$

where

(1.10) $$Z_V(y)=\int\exp(-U_V(x)-I_{V,V^c}(xy))\,d\omega_V(x).$$

All probability measures ν on X for which the above conditional distributions are a version of $E(.\,|\,\mathscr{B}(V^c))$ are called *equilibria*. If there is more than one equilibrium, we say that phase transition occurs. The set of all stationary tempered equilibria will be denoted by $\mathscr{G}_0(U)$. A result of Lebowitz-Presutti ([5], Theorems 4.3 and 4.5) tells us that *$\mathscr{G}_0(U)$ is not empty when U is superstable and regular.* The proof of this result uses the following basic estimate of Ruelle [12]:

(1.11) **Theorem.** *If the energy U is superstable and regular, then there exists to any $\gamma < A$ a δ such that for all $V \subseteq W \in \mathscr{V}$:*

$$\int \pi^W(x \mid 0)\, d\omega_{W \setminus V}(x_{W \setminus V}) \leqq \exp\Big(-\gamma \sum_{i \in V} x_i^2 + \delta |V|\Big)$$

(0 is the boundary condition y with $y_i = 0$ for all $i \notin W$).

The same arguments as in the proof of this theorem show that *the same estimate holds also for $\nu_V(x)$, the density of $\nu|_{\mathscr{B}(V)}$ with respect to $d\omega_V(x)$ if ν is in $\mathscr{G}_0(U)$ with U superstable and regular* (see Lebowitz-Presutti [5], Theorem 4.4, or Ruelle [11]).

We will use these estimates in Sect. 3.

(1.12) *Examples.* Let $n = 1$.

i) $U_V(x) = \sum_{i \in V} x_i^2 + \sum_{i \neq j} a_{i-j} x_i x_j$ with $a_k = a_{-k}$ and $\sum_k |a_k| < \infty$. Then regularity is satisfied under a mild condition on the decay of the a_k and implies also superregularity. We introduce $P(x) = 1 - \sum_{k \neq 0} a_k e^{ikx}$. It is easily seen that U is superstable iff $\inf_x P(x) > 0$, in fact $A = \inf_x P(x)$, $B = 0$. If $d\omega(x)$ is the Lebesgue-measure, then it can be shown by generalizing the results of Rosanov [10], Spitzer [13] and Künsch [3] that the Gaussian field with spectral density $(2\pi)^{-d} \frac{1}{2} P(x)^{-1}$ is in $\mathscr{G}_0(U)$. But in order to have $\mathscr{G}_0(U) \neq \emptyset$ it is sufficient to have $P(x) \geqq 0$ and $\int_{[-\pi,\pi]^d} P(x)^{-1}\, dx < \infty$. So we have examples where the superstability condition is not satisfied.

ii) $U_V(x) = \sum_{i \in V} F(x_i) + \sum_{i \neq j \neq k} b_{ijk} G(x_i, x_j, x_k)$ where b_{ijk} is invariant under translations and permutations of the indices and G is symmetric in its arguments. If $\sum_{j \neq k} |b_{ijk}| < \infty$ and $|G(x, y, z)| \leqq \text{const.}(x^2 + y^2 + z^2)$, then this energy is regular, and if F is bounded below and grows quick enough at infinity, it is also superstable. Like this example we can construct energies where the potential Φ_V is not zero for arbitrarily big $|V|$.

2. Energy and Mean Entropy

We want to define the average energy of a configuration x as the limit of $|V|^{-1} U_V(x)$ as V tends to $\mathbb{Z}^d$. We will prove the existence of this limit by using ergodic theorems for $\mathbb{Z}^d$. It is known that for these theorems it is necessary that V tends to $\mathbb{Z}^d$ in some regular way which we are going to define now.

(2.1) *Definition.* For $F \in \mathscr{V}$ we define the "*r-inside*" $\mathring{V}(r)$ as the set $\{i \in V, |i - j| > r$ for all $j \notin V\}$ and the "*r-outside*" $\bar{V}(r)$ as the set $\{i \in \mathbb{Z}^d, |i - j| \leqq r$ for a $j \in V\}$.

(2.2) *Definition.* A sequence (V_n) is called a *van Hove* sequence if $V_n \in \mathscr{V}$, V_n increases monotonically to $\mathbb{Z}^d$ and for all $r \in \mathbb{N}$ $\lim |\bar{V}_n(r) \setminus \mathring{V}_n(r)| / |V_n| = 0$. It is called *regular* if moreover there exists a sequence of cubes $F_n \supseteq V_n$ such that $\sup |F_n| / |V_n| < \infty$.

A van Hove sequence allows us to neglect boundary terms, and a regular sequence is also "round enough". If now Z is a random variable and μ is stationary with $E_\mu(|Z|)<\infty$, then the following *ergodic theorem* holds (see Tempel'man [14] or Nguyen and Zessin [6]): $|V_n|^{-1}\sum_{i\in V_n} Z\circ\theta_i \to E(Z\,|\,\mathscr{I})$ ($\mathscr{I}$ is the σ-field of invariant events) in $L_1(d\mu)$ if (V_n) is van Hove and a.s. $(d\mu)$ if (V_n) is regular.

Except in trivial cases the energy $U_V(x)$ is not of the form $\sum_{i\in V} Z\circ\theta_i$, but the regularity condition allows us to control the difference between $U_V(x)$ and an additive expression. We have

(2.3) **Lemma.** *Let $V=\bigcup_{k=1}^{m} V_k$ with $V_i\cap V_j=\emptyset$ $(i\neq j)$ and let U be regular. Then we have*

$$\left|U_V-\sum_{k=1}^{m} U_{V_k}\right|\leqq\frac{1}{2}\sum_{k=1}^{m}\sum_{i\in V_k} x_i^2\Big(\sum_{j\notin V_k}\psi(|i-j|)\Big).$$

Proof. By induction on m.

Let us sketch first how one can reduce the convergence of $|V_n|^{-1}U_{V_n}(x)$ to the additive case by using the above lemma: we partition $\mathbb{Z}^d$ in cubes of size p^d parallel to the axes:

$$F_{p,k}=\{j\in\mathbb{Z}^d, (k_i-1)\,p<j_i\leqq k_i p \quad (i=1,\ldots,d)\} \quad (k\in\mathbb{Z}^d, p\in\mathbb{N}).$$

Let $J_{n,p}$ be the following set $\{k\in\mathbb{Z}^d, F_{p,k}\subseteq V_n\}$. For any $p\in\mathbb{N}$ and $q<p/2$ V_n is then the union of three disjoint subsets: $V_n^1=\bigcup_{k\in J_{n,p}}\mathring{F}_{p,k}(q)$, $V_n^2=\bigcup_{k\in J_{n,p}}(F_{p,k}\setminus\mathring{F}_{p,k}(q))$ and $V_n^3=V_n\setminus V_n^1\setminus V_n^2$. We write now $U_{V_n}=\sum_{k\in J_{n,p}} U_{F_{p,k}}+$error, and because of Lemma (2.3) the error is bounded by $\sum_{i\in V_n} w_i X_i^2+\sum_{i\in V_n^3}|U_i(x_i)|$ with some weights w_i. If q is big, then the w_i are small for all $i\in V_n^1$, and if p and n are also big, most points of V_n will belong to V_n^1.

We are going now to make the above arguments exact.

(2.4) **Lemma.** *Let μ be stationary with $E_\mu(X_i^2)<\infty$. For all p,q $\lim|V_n|^{-1}\sum_{i\in V_n^2} X_i^2$ exists in $L_1(d\mu)$ if (V_n) is van Hove and a.s. $(d\mu)$ if (V_n) is regular. For fixed q these limits tend to zero in $L_1(d\mu)$ for $p\to\infty$.*

Proof. We have $|V_n|^{-1}\sum_{i\in V_n^2} X_i^2=p^d|J_{n,p}|/|V_n|\cdot|J_{n,p}|^{-1}\sum_{k\in J_{n,p}} Z_k$ where $Z_k=p^{-d}\cdot\sum_{i\in F_{p,k}\setminus\mathring{F}_{p,k}(q)} X_i^2$. The first ratio on the right tends to 1 because of the van Hove condition, and the second term converges because of the ergodic theorem since $(J_{n,p})_{n\in\mathbb{N}}$ is also van Hove and regular if (V_n) is so as one sees easily. Furthermore because of L_1-convergence $E(\lim|V_n|^{-1}\sum_{i\in V_n^2} X_i^2)=E(X_i^2)(1-(1-2q/p)^d)$. so the last claim follows also. □

(2.5) **Theorem.** *Let U be a stationary and regular energy and μ stationary with $E_\mu(X_i^2)<\infty$ and $E_\mu|U_i|<\infty$. $|V_n|^{-1}U_{V_n}(x)$ converges then in $L_1(d\mu)$ for any* *van Hove sequence and also a.s.* $(d\mu)$ *if* (V_n) *is regular. The limit does not depend on the sequence* (V_n).

Proof. We are going to show a.s. convergence (the arguments for L_1-convergence are the same). Fix a regular sequence (V_n). We want to show that $|V_n|^{-1} U_{V_n}(x)$ is a.s. a Cauchy sequence.

Let $\varepsilon>0$ be given. First, we choose q such that $\sum_{|k|>q} \psi(|k|)<\varepsilon$, and after a p such that $\lim_n |V_n|^{-1} \sum_{i\in V_n^2} X_i^2<\varepsilon$ (this is possible because of Lemma (2.4): By choosing a subsequence we get also a.s. convergence in p). It then follows from Lemma (2.3):

$$(2.6)\qquad \begin{aligned} &||V_n|^{-1} U_{V_n}(x)-|V_m|^{-1} U_{V_m}(x)| \\ &\leqq ||V_n|^{-1} \sum_{k\in J_{n,p}} U_{F_{p,k}} - |V_m|^{-1} \sum_{k\in J_{m,p}} U_{F_{p,k}}| \\ &\quad +\sum_k \psi(|k|)\,|V_n|^{-1} \sum_{i\in V_n^3} X_i^2 + |V_n|^{-1} \sum_{i\in V_n^3} |U_i(x)| \\ &\quad + \sum_{|k|>q} \psi(|k|)\,|V_n|^{-1} \sum_{i\in V_n^1} X_i^2 + \sum_k \psi(|k|)\,|V_n|^{-1} \sum_{i\in V_n^2} X_i^2 \\ &\quad + \text{same terms with } m \text{ instead of } n. \end{aligned}$$

Because of the van Hove condition $|V_n|\sim p^d\cdot|J_{n,p}|$, and because of the ergodic theorem $|J_{n,p}|^{-1} \sum_{k\in J_{n,p}} U_{F_{p,k}}$ is convergent. Therefore the first term on the right hand side of (2.6) is $<\varepsilon$ if n and m are big enough.

Since $|V_n|\sim|V_n^1\cup V_n^2|$, $\lim|V_n|^{-1} \sum_{i\in V_n} X_i^2 = \lim|V_n|^{-1} \sum_{i\in V_n^1\cup V_n^2} X_i^2$. Therefore, the second term is also $<\varepsilon$ for n big enough, and the same argument applies also to the third term. Combining these results:

$$(2.7)\qquad \begin{aligned} &||V_n|^{-1} U_{V_n}(x)-|V_m|^{-1} U_{V_m}(x)| \\ &\leqq 5\varepsilon+2\varepsilon\sum_k \psi(|k|)+2\varepsilon \sup_n |V_n|^{-1} \sum_{i\in V_n} X_i^2 \end{aligned}$$

for n and m big enough. The last supremum is finite because of the ergodic theorem.

Using L_1-convergence and the same arguments as before, we find:

$$\begin{aligned} &E_\mu|\lim|V_n|^{-1} U_{V_n}(x)-\lim|V_n|^{-1} \sum_{k\in J_{n,p}} U_{F_{p,k}}(x)| \\ &\leqq E_\mu(X_i^2) \sum_{|k|>q} \psi(|k|)+E_\mu(X_i^2)\sum_k \psi(|k|)(1-(1-2q/p)^d) \end{aligned}$$

which is arbitrarily small if p and q are big enough. But

$$\lim|V_n|^{-1} \sum_{k\in J_{n,p}} U_{F_{p,k}}(x)=p^{-d} E(U_{F_{p,0}}|\mathscr{I}_p)$$

where $\mathscr{I}_p$ is the σ-field invariant unter translations pk, $k\in\mathbb{Z}^d$. Therefore $\lim_n |V_n|^{-1} U_{V_n}(x)=\lim_p p^{-d} E(U_{F_{p,0}}|\mathscr{I}_p)$ independent of (V_n). □

(2.8) *Definition.* The *mean energy e* of a stationary field μ is defined as

$e(\mu)=\lim |V_n|^{-1} E_\mu(U_{V_n}(x))$ if $E_\mu|U_V|<\infty$ $(V\in\mathscr{V})$ and the limit exists for all van Hove sequences and is independent of the chosen sequence,

$=+\infty$ otherwise.

(2.9) *Remarks.* i) Usually the convergence of energy is proved under a condition on the potential Φ, e.g. we have L_1- and a.s. convergence if $\sum_{V\ni 0} |V|^{-1}|\Phi_V|$ is in L_1 (Föllmer, private communication, see also Nguyen-Zessin [6], Remark (7.21)). But if $\Phi_V \neq 0$ for $|V|>2$ it is not clear whether the above condition follows from regularity. From (1.4) we get for instance $\Phi_{\{i,j,k\}} = I_{\{i\},\{j,k\}} - I_{\{i\},\{j\}} - I_{\{i\},\{k\}}$. If we use regularity for every term, we get $+\infty$ after summation over all j and k.

ii) In the proof of Theorem (2.5) we have never used the consistency conditions (1.2). Moreover, the same arguments apply if we replace X_i^2 in the definition of regular by any $Z\circ\theta_i$. However, we do not know if these ergodic theorems follow also from the results of Nguyen-Zessin [6]. We notice also that we did not show in Theorem (2.5) that the limit is $\mathscr{I}$-measurable.

If μ is stationary with $\mu|_{\mathscr{B}(V)} \ll d\omega_V$, we denote the density of $\mu|_{\mathscr{B}(V)}$ with respect to $d\omega_V$ by $\mu_V(x)$. As in the case of the energy we would like to define the average entropy as the limit of $-|V|^{-1}\log\mu_V(x)$, but at the present stage we are not able to prove the convergence of this expression. It will follow easily after the estimates of Sect. 3. At the moment we recall a result of Ruelle which says that at least $|V|^{-1}E_\mu(\log\mu_V(x))$ converges.

(2.10) **Theorem** (Ruelle). *If μ is stationary with $E_\mu(X_i^2)<\infty$ and $\mu|_{\mathscr{B}(V)}\ll d\omega_V$, then for every van Hove sequence (V_n) $|V_n|^{-1}E_\mu(\log\mu_V(x))$ converges to $\sup_p |F_p|^{-1}E_\mu(\log\mu_{F_p}(x))$ ($+\infty$ as value for* lim *and* sup *must be allowed).*

Proof. If $d\omega$ is a probability measure on $\mathbb{R}^n$, the result is well known, see for instance Preston [8], Theorem 8.1. If $d\omega$ is not a probability measure, we consider $d\tilde{\omega}(x)=\exp(a-x^2)\,d\omega(x)$ with a suitable a. □

(2.11) *Definition.* The *mean entropy* of a stationary field μ is defined as

$s(\mu)=-\lim|V_n|^{-1}E_\mu(\log\mu_{V_n}(x))$ if the density $\mu_V(x)$ exists and is integrable and if the limit exists for every van Hove sequence and is independent of the chosen sequence,

$=-\infty$ otherwise.

3. Absolute and Conditional Density

If ν is an equilibrium then $\nu|_{\mathscr{B}(V)}\ll d\omega_V$, and the density is given by

(3.1) $$\nu_V(x)=\int \pi^V(x\,|\,y)\,d\nu_{V^c}(y).$$

Since the interaction between distant points is small, it is intuitively clear that the different boundary conditions y will not have much influence in the interior of V. This would mean that we can approximate the right hand side of (3.1) by $\pi^V(x\,|\,y_0)$ with some fixed boundary condition y_0. For simplicity we take $y_0=0$. We will give in this section a precise statement in which sense $\nu_V(x)$ and $\pi^V(x\,|\,0)$ are very close for big V. First we prove a couple of lemmas which all follow from Jensen's inequality.

(3.2) **Lemma** (Pirlot [7]). *If* $\mu|_{\mathscr{B}(V)} \ll d\omega_V$ *and* $U_V(x)$ *and* $\log \mu_V(x)$ *are integrable, then* $E_\mu(U_V(x)) + E_\mu(\log \mu_V(x)) \geqq -\log Z_V(0)$.

Proof.

$$\int (\log(\exp(-U_V(x))) - \log \mu_V(x))\, \mu_V(x)\, d\omega_V(x)$$
$$\leqq \log \int \exp(-U_V(x))\, \mu_V(x)^{-1}\, \mu_V(x)\, d\omega_V(x) = \log Z_V(0). \quad \square$$

(3.3) **Lemma** (Pirlot [7]). *Suppose U is regular and superstable. Then*

$$\log Z_V(y) \geqq \log Z_V(0) - \tfrac{1}{2} \sum_{i\in V,\, j\notin V} \psi(|i-j|)(K + y_j^2).$$

Proof.

$$\log(Z_V(y)/Z_V(0)) = \log \int \exp(-I_{V,V^c}(x\,y))\, \pi^V(x\,|\,0)\, d\omega_V(x)$$
$$\geqq -\int I_{V,V^c}(x\,y)\, \pi^V(x\,|\,0)\, d\omega_V(x)$$
$$\geqq -\tfrac{1}{2} \sum_{i\in V,\, j\notin V} \psi(|i-j|)\left(\int x_i^2\, \pi^V(x\,|\,0)\, d\omega_V(x) + y_j^2\right).$$

Now we apply the estimates of Theorem (1.11). $\square$

(3.4) **Lemma** (Pirlot [7]). *If U is regular and superstable and if ν is in $\mathscr{G}_0(U)$, then we have*

$$E_\nu(\log \nu_V(x)) \leqq -E_\nu(\log Z_V(x)) - E_\nu(U_V(x)) + K \sum_{i\in V,\, j\notin V} \psi(|i-j|).$$

Proof. By (3.1) and Jensen's inequality for the function $x \log(x)$:

$$\log \nu_V(x)\, \nu_V(x) \leqq \int \pi^V(x\,|\,y) \log(\pi^V(x\,|\,y))\, d\nu_{V^c}(y)$$
$$\leqq \int \pi^V(x|y)\left(-\log Z_V(y) - U_V(x) + \tfrac{1}{2} \sum_{i\in V,\, j\notin V} \psi(|i-j|)(x_i^2 + y_j^2)\right) d\nu_{V^c}(y).$$

Now because of Theorem (1.11) $E_\nu(X_i^2) < \infty$. Superstability shows that $U_V(x)^-$ is in $L_1(d\nu)$, and finally Lemma (3.3) shows that $(\log Z_V(x))^-$ is also in $L_1(d\nu)$. Therefore the claim of the lemma follows by integrating both sides of the above inequality with respect to $d\omega_V(x)$ and applying Fubini. $\square$

From these inequalities we get

(3.5) **Corollary.** *If U is regular and superstable and ν is in $\mathscr{G}_0(U)$, then*

i) *U_V is integrable, and*

ii) $E_\nu(\log Z_V(x)) \leqq \log Z_V(0) + K \sum_{i\in V,\, j\notin V} \psi(|i-j|)$.

Proof. i) follows from (3.4): Since $U_V(x)^-$ and $(\log Z_V(x))^-$ are in $L_1(d\nu)$, the right hand side of (3.4) can only be equal to $-\infty$. But the left hand side is bounded below: If $d\omega$ is a probability measure, this is obvious, otherwise we change $d\omega$ to a probability measure $d\tilde{\omega}$ as in the proof of Theorem (2.10) and use that $E_\nu(X_i^2)<\infty$.

ii) is just a combination of the Lemmas (3.2) and (3.4). □

We are going now to estimate the difference $\log \nu_V(x)-\log \pi^V(x\,|\,0)$ for $\nu\in\mathscr{G}_0(U)$. We want to show that this difference is of smaller order than $|V|$ for many configurations x. By definition and (3.1)

$$(3.6)\qquad \begin{aligned}&\log \nu_V(x)-\log \pi^V(x\,|\,0)\\ &\quad=\log\left(\int \exp\left(-I_{V,V^c}(x\,y)\right) Z_V(0)/Z_V(y)\, d\nu_{V^c}(y)\right).\end{aligned}$$

A lower bound follows directly from Jensen's inequality:

(3.7) **Lemma.** *Suppose U is regular and superstable and ν is in $\mathscr{G}_0(U)$. Then there exists K independent of V such that*

$$\log \nu_V(x)-\log \pi^V(x\,|\,0)\geqq -\tfrac{1}{2}\sum_{i\in V,\, j\notin V} \psi(|i-j|)(x_i^2+K).$$

Proof. Applying Jensen's inequality to (3.6) we get

$$\begin{aligned}&\log \nu_V(x)-\log \pi^V(x\,|\,0)\\ &\quad\geqq -\int I_{V,V^c}(x\,y)\, d\nu_{V^c}(y)+\log Z_V(0)-E_\nu(\log Z_V(x)).\end{aligned}$$

Now we use regularity, Theorem (1.11) and Corollary (3.5) ii). □

The upper bound is much more delicate.

(3.8) **Lemma.** *Suppose U is superregular and superstable, and ν is in $\mathscr{G}_0(U)$. Then there exist K_1, K_2, K_3, K_4 independent of V such that*

$$\begin{aligned}\log \nu_V(x)-\log \pi^V(x\,|\,0)&\leqq K_1 \sum_{i\in V,\, j\notin V} \psi(|i-j|)(|x_i|+K_2)\\ &\quad+K_3\sum_{j\notin V}\Big(\sum_{i\in V}\psi(|i-j|)(|x_i|+K_2)\Big)^2+K_4.\end{aligned}$$

Proof. If we have superregularity instead of regularity, we can improve the estimate of Lemma (3.3):

$$\log Z_V(y)-\log Z_V(0)\geqq -K\sum_{i\in V,\, j\notin V}\psi(|i-j|)\,|y_j|.$$

Because of (3.6) it is therefore sufficient to give an estimate of $E_\nu(\exp\sum_{j\notin V} a_j|X_j|)$ where $a_j=\sum_{i\in V}(|x_i|+K)\,\psi(|i-j|)$ for short. Put $a=\sum_{j\notin V} a_j$. Using integration by parts we find for any $N>0$:

$$(3.9)\qquad \begin{aligned}E_\nu\Big(\exp\sum_{j\notin V} a_j|X_j|\Big)&\leqq e^{aN}\Big(1+\sum_{k=1}^{\infty} e^k\, \nu\Big[aN+k-1<\sum_{j\notin V} a_j|X_j|\leqq aN+k\Big]\Big)\\ &=e^{aN}\Big(1+e+(e-1)\sum_{k=1}^{\infty} e^k\, \nu\Big[\sum_{j\notin V} a_j|X_j|>aN+k\Big]\Big).\end{aligned}$$

We are now going to find a bound for $\nu[\sum_{j\notin V} a_j |X_j| > aN+k]$. First we notice that Theorem (1.11) implies that

$$\text{(3.10)} \qquad \nu[\sum_{i\in V} X_i^2 \geqq c] \leqq \exp(\delta|V| - \bar{\gamma}c) \int \exp(-(\gamma-\bar{\gamma})\sum_{i\in V} x_i^2)\, d\omega_V(x)$$
$$= \exp(\bar{\delta}|V| - \bar{\gamma}c) \qquad \text{for all } V \text{ and } c \text{ and } \bar{\gamma} < \gamma < A.$$

So by Schwartz inequality

$$\text{(3.11)} \qquad \nu[\sum_{j\in V} |X_j| \geqq |V| N] \leqq \nu[\sum_{j\in V} X_j^2 \geqq |V| N^2] \leqq \exp(-(\bar{\gamma}N^2 - \bar{\delta})|V|).$$

Now take an enumeration $(j_n)_{n\in\mathbb{N}}$ of V^c such that $a_{j_1} \geqq a_{j_2} \geqq \dots$ and put $W_n = \{j_1, j_2, \dots, j_n\}$. Then we get by integration by parts

$$\sum_{j\notin V} a_j |X_j| = \lim_m \left(\sum_{i=1}^{m} (a_{j_i} - a_{j_{i+1}}) \sum_{k\in W_i} |X_k| + a_{j_{m+1}} \sum_{k\in W_m} |X_k| \right).$$

(3.11) and the Lemma of Borel-Cantelli show that a.s. $\sum_{j\in W_n} |X_j| > nN$ only for finitely many n if we choose $N^2 > \bar{\delta}/\bar{\gamma}$. If $\sum_{k\in W_n} |X_k| \leqq nN$ for all $n > r$ and $\sum_{k\in W_r} |X_k| > rN$, then we have

$$\begin{aligned}
\sum_{j\notin V} a_j |X_j| - aN &= \lim_m \left(\sum_{i=1}^{m} (a_{j_i} - a_{j_{i+1}}) (\sum_{k\in W_i} |X_k| - iN) + a_{j_{m+1}} (\sum_{k\in W_m} |X_k| - mN) \right) \\
&\leqq \sum_{i=1}^{r} (a_{j_i} - a_{j_{i+1}}) (\sum_{k\in W_i} |X_k| - iN) \\
&= \sum_{i=1}^{r} a_{j_i} |X_{j_i}| - N \sum_{i=1}^{r} a_{j_i} - a_{j_{r+1}} (\sum_{k\in W_r} |X_k| - rN) \\
&\leqq \sum_{j\in W_r} a_j |X_j|.
\end{aligned}$$

This shows that $\sum_{j\in W_r} a_j |X_j|$ must be $> k$ if $\sum_{j\notin V} a_j |X_j| > aN + k$ and $\sum_{j\in W_n} |X_j| \leqq nN$ for all $n > r$ but $\sum_{j\in W_r} |X_j| > rN$. But this means that we can replace the set $[\sum_{j\notin V} a_j |X_j| > aN + k]$ which depends on infinitely many coordinates by a union of sets which all depend only on finitely many coordinates and whose probability can be estimated using (3.11). Namely we have

$$\begin{aligned}
\text{(3.12)} \qquad \nu[\sum_{j\notin V} a_j |X_j| > aN + k] &\leqq \sum_{r=1}^{\infty} \nu[\sum_{j\in W_r} a_j |X_j| > k, \sum_{j\in W_r} |X_j| > rN] \\
&\leqq \sum_{r=1}^{\infty} \nu[\sum_{j\in W_r} X_j^2 > k^2 / \sum_{j\in W_r} a_j^2, \sum_{j\in W_r} |X_j| > rN] \\
&\leqq \sum_{r=1}^{\infty} \exp(\bar{\delta} r - \max(rN^2, k^2/\sum_{j\notin V} a_j^2)\, \bar{\gamma}).
\end{aligned}$$

Putting $\bar{a}=\sum_{j\notin V} a_j^2$ and inserting (3.12) in (3.9) we find

$$\sum_{k=1}^{\infty} e^k\, v[\sum_{j\notin V} a_j|X_j|>aN+k] \leqq \sum_{r=1}^{\infty} \exp(r(\bar{\delta}-\bar{\gamma}N^2)) \sum_{k=1}^{N(\bar{a}r)^{1/2}} e^k$$
$$+\sum_{k=1}^{\infty} \exp(k-\bar{\gamma}k^2/\bar{a}) \sum_{r=1}^{k^2/(N^2\bar{a})} e^{\bar{\delta}r}$$
$$\leqq \frac{e}{e-1} \sum_{r=1}^{\infty} \exp(r(\bar{\delta}-\bar{\gamma}N^2)+N(\bar{a}r)^{1/2})$$
$$+e^{\bar{\delta}}/(e^{\bar{\delta}}-1) \sum_{k=1}^{\infty} \exp(k-k^2(\bar{\gamma}N^2-\bar{\delta})/(N^2\bar{a})).$$

Using Lemma (3.13) below we get

$$E_v(\exp \sum_{j\notin V} a_j|X_j|) \leqq e^{aN}(1+e+\exp(N^2\bar{a}/4(\bar{\gamma}N^2-\bar{\delta}))(K_1+K_2\bar{a}^{1/2})).$$

In order to complete the proof we take logarithms using that $\log(a+b) \leqq \log a + \log b$ for a and $b \geqq 2$ and go back to the definitions of a and $\bar{a}$. □

Finally we give a lemma announced in the proof before.

(3.13) **Lemma.** *There exist constants K_1, K_2 such that for all a and $b>0$:*

$$\sum_{k=1}^{\infty} \exp(k-ak^2) \leqq \exp(1/4a)(1+K_1 a^{-1/2})$$

and

$$\sum_{k=1}^{\infty} \exp(bk^{1/2}-ak) \leqq \exp(b^2/4a)(1+K_1 a^{-3/2} b + K_2 a^{-1}).$$

Proof. Because $\exp(x-ax^2)$ is increasing for $x<(2a)^{-1}$ and decreasing for $x>(2a)^{-1}$ we have:

$$\sum_{k=1}^{\infty} (\exp(k-ak^2)) \leqq \int_1^{\infty} \exp(x-ax^2)\,dx + \exp(1/4a)$$
$$=\exp(1/4a)(1+\int \exp(-a(x-1/2a)^2)\,dx) \leqq \exp(1/4a)(1+a^{-1/2}K)$$

where $K=\int_{-\infty}^{+\infty} \exp(-x^2)\,dx$. The proof of the second claim is analogous. □

Now we are able to prove our central result.

(3.14) **Theorem.** *Let U be superregular and superstable, v in $\mathscr{G}_0(U)$ and μ stationary with $E_\mu(X_i^2)<\infty$. Then $|V_n|^{-1}(\log v_{V_n}(x)-\log \pi^{V_n}(x\,|\,0))$ converges to zero in $L_1(d\mu)$ for every van Hove sequence and a.s. $(d\mu)$ if (V_n) is regular.*

Proof. We only have to show that the bounds of Lemma (3.7) and (3.8) go to zero after division by $|V|$. All terms which occur can be treated the same way, therefore we give the proof only for $\sum_{j\notin V}(\sum_{i\in V}|X_i|\psi(|i-j|))^2$. We put

$Y_i = \sum_{j \in \mathbb{Z}^d} \psi(|i-j|)\,|X_j|$. Then we have

$$\sum_{j \notin V} \sum_{i \in V} |X_i|\,\psi(|i-j|) \sum_{k \in V} |X_k|\,\psi(|k-j|) \leqq \sum_{i \in V,\, j \notin V} |X_i|\,Y_j\,\psi(|i-j|).$$

If we call the last expression T_V, we can prove the convergence of $|V|^{-1} T_V$ as in Theorem (2.5) (cf. Remark (2.9) ii)) since we have

$$T_{V \cup W} - T_V - T_W = - \sum_{i \in V,\, j \in W} \psi(|i-j|)(|X_i|\,Y_j + |X_j|\,Y_i)$$

which is bounded by $\sum_{i \in V,\, j \in W} \psi(|i-j|)(X_i^2 + Y_i^2 + X_j^2 + Y_j^2)$. Furthermore $E(Y_i^2) \leqq (\sum_i \psi(|i|))^2 E(X_i^2)$. In order to see that the limit is zero we notice that

$$E_\mu(T_V) \leqq \sum_{i \in V,\, j \notin V} \psi(|i-j|) \sum_k \psi(|k|)\, E_\mu(X_i^2)$$

which is of smaller order than $|V|$. □

The next theorem treats the case where we have only regularity instead of superregularity.

(3.15) **Theorem.** *Theorem* (3.14) *holds also for regular superstable energies U for which* $\sum_k \psi(|k|) < A$, *A being the constant in the definition of superstable.*

Proof. We give only those arguments which differ from those used in (3.8) and (3.14). For a corresponding result as Lemma (3.8) it is sufficient to find an upper bound of $E_\nu(\exp \sum_{j \notin V} b_j X_j^2)$ where $b_j = \sum_{i \in V} \psi(|i-j|)$. We put $b = \sum_{j \notin V} b_j$. Then (3.9) carries over without problems, and with the same arguments which led to (3.12) we get

$$\text{(3.16)} \qquad \nu[\sum_{j \notin V} b_j X_j^2 > bN + k] \leqq \sum_{r=1}^{\infty} \nu[\sum_{j \in W_r} b_j X_j^2 > k,\ \sum_{j \in W_r} X_j^2 > rN]$$

$$\leqq \sum_{r=1}^{\infty} \min(\nu[\sum_{j \in W_r} b_j X_j^2 > k],\ \nu[\sum_{j \in W_r} X_j^2 > rN]).$$

Here we have now to use a different argument in order to estimate $\nu[\sum_{j \in W_r} b_j X_j^2 > k]$ because Schwartz inequality does not help any more. Because of the assumption $\sum_k \psi(|k|) < A$ we can choose $\gamma < A$ and $\eta > 1$ such that $\gamma - \eta \sum_k \psi(|k|) > 0$. Now it is not difficult to see that because of the condition on $d\omega$ there exists a constant M such that for all c with $\gamma > c > \gamma - \eta \sum_k \psi(|k|)$ we have

$$\int \exp(-cx^2)\, d\omega(x) \leqq \exp(M(\gamma - c)) \cdot \int \exp(-\gamma x^2)\, d\omega(x).$$

But this implies that

$$\nu[\sum_{j \in W_r} b_j X_j^2 > k] \leqq \exp(\delta r - \eta k) \int \exp(-\sum_{j \in W_r} (\gamma - \eta b_j)\, x_j^2)\, d\omega_{W_r}(x)$$

$$\leqq \exp(\bar{\delta} r - \eta(k - M \sum_{j \in W_r} b_j)) \leqq \exp(\bar{\delta} r - \eta(k - Mb)).$$

So we have found

$$(3.17)\qquad \nu\Big[\sum_{j\in W_r} b_j X_j^2 > bN+k\Big] \leqq \sum_{r=1}^{\infty} \exp(\bar{\delta} r - \max(\bar{\gamma} N r, \eta(k-Mb)))$$

which can be used instead of (3.12). The rest of the proof carries over with minor changes. □

(3.18) *Remark.* If we want to compare $\log \nu_V(x)$ and $\log \pi^V(x\,|\,y)$ for an arbitrary but fixed boundary condition y instead of 0 we need to estimate the difference $\log Z_V(y) - \log Z_V(0)$ since it can be seen immediately that $|V|^{-1} I_{V,V^c}(xy)$ converges to zero. This would require again rather complicated estimates, see for instance Lebowitz-Presutti [5], Chap. 3. We leave this problem because we can deduce all desired results already from the Theorems (3.14) and (3.15).

4. Variational Principle, almost Sure Entropy and Information Gain

We start with the following identity

$$(4.1)\qquad \log \nu_V(x) = -\log Z_V(0) - U_V(x) + (\log \nu_V(x) - \log \pi^V(x\,|\,0)).$$

From this we first get a new proof of results of Lebowitz-Presutti [5] (existence of pressure) and Pirlot [7] (first part of the variational principle).

(4.2) **Theorem.** *If U is stationary, superstable and superregular (or regular with $\sum \psi(|k|) < A$) and if ν is in $\mathscr{G}_0(U)$, then the limit of $|V_n|^{-1} \log Z_{V_n}(0)$ exists for any van Hove sequence and is equal to $-(e(\nu)-s(\nu)) = -\inf\{e(\mu)-s(\mu),\ \mu$ stationary$\}$ independently of the chosen sequence.*

The limit of $|V_n|^{-1} \log Z_{V_n}(0)$ is called the *pressure* and is denoted by p. The difference $e(.)-s(.)$ is called the *free energy*.

Proof. We divide (4.1) by $|V|$ and take expectation with respect to ν. Because of Theorem (1.11) $E_\nu(X_i^2) < \infty$, and because of Corollary (3.5)i) also $E_\nu |U_V(x)| < \infty$. The existence of the limit and the identity $p = -(e(\nu)-s(\nu))$ follow then from the Theorems (2.5), (2.10) and (3.14) respectively (3.15). Finally Lemma (3.2) implies the variational inequality $-p \leqq e(\mu) - s(\mu)$. □

Together with the previous results (4.1) gives also the almost sure convergence of entropy.

(4.3) **Theorem.** *Suppose ν is in $\mathscr{G}_0(U)$ with a superstable and superregular energy U (or regular with $\sum_k \psi(|k|) < A$). Then $|V_n|^{-1} \log \nu_{V_n}(x)$ converges in L_1 for every van Hove sequence (V_n) and also a.s. $(d\nu)$ if (V_n) is a regular sequence, and the limit is equal to $-p - \lim |V_n|^{-1} U_{V_n}(x)$.*

Proof. Divide (4.1) by $|V|$ and take the limit with respect to V without taking expectations. Then all terms on the right hand side converge, due to Theorems (2.5), (4.2) and (3.14) respectively (3.15). □

(4.4) *Remark.* L_1-convergence, but not the a.s. convergence can be proved under weaker conditions with the methods of Nguyen-Zessin [6]. The idea to prove the a.s. convergence for Gibbs states via the a.s. convergence of the energy and the pressure is due to Föllmer [1].

Theorem (4.3) says that we can calculate the entropy of an equilibrium from one typical configuration and with the energy U alone without knowing which ν out of $\mathscr{G}_0(U)$ has been realized. Moreover we can get also an almost sure version of the variational principle.

(4.5) **Corollary.** *Let $U^{(i)}$, $i=1,2$, be two superstable and superregular energies (or regular with $\sum_k \psi(|k|)<A$) with potentials $\Phi^{(i)}$, and let ν be in $\mathscr{G}_0(U^{(1)})$ such that $\sum_{V\ni 0} |\Phi_V^{(i)}||V|^{-1}\in L_1(d\nu)$ for $i=1,2$. Then we have for every regular sequence (V_n) a.s. $(d\nu)$:*

$$\lim |V_n|^{-1}(U_{V_n}^{(2)}(x)+\log \nu_{V_n}(x))\geqq -p^{(2)}, \; = -p^{(2)} \quad \text{if } U^{(1)}=U^{(2)}.$$

Proof. Under the above conditions $\lim |V_n|^{-1} U_{V_n}^{(i)}(x)$ is $\mathscr{I}$-measurable (cf. Remark (2.9)). Therefore the corollary follows from the Theorems (4.2) and (4.3) if ν is ergodic. But every stationary equilibrium is a mixture of ergodic equilibria, see Föllmer [2]. □

From (4.1) we can deduce a third result, but for this we first have to define the information gain.

(4.6) *Definition.* Let ν and μ be stationary. The *information gain* of μ with respect to ν is defined as

$$h(\mu,\nu)=\lim |V|^{-1} H_V(\mu,\nu), \text{ where } H_V(\mu,\nu)=E_\mu\left(\log \frac{d\mu|_{\mathscr{B}(V)}}{d\nu|_{\mathscr{B}(V)}}(x)\right),$$

if $\frac{d\mu|_{\mathscr{B}(V)}}{d\nu|_{\mathscr{B}(V)}}$ and the above integral exist and if the limit exists for all van Hove sequences (independently of the chosen sequence).

$= +\infty$ otherwise.

Using Jensen's inequality it follows easily that $H_V(\mu,\nu)=0$ iff $\mu|_{\mathscr{B}(V)}=\nu|_{\mathscr{B}(V)}$. Because of the division by $|V|$ and the passage to the limit this need not to be true for $h(\mu,\nu)$ any more, in fact we will have counterexamples below. Let us first show the connection between the information gain and the quantities defined before.

(4.7) **Theorem.** *Let μ be stationary with $E_\mu(X_i^2)<\infty$ and $s(\mu)>-\infty$, and let ν be in $\mathscr{G}_0(U)$ for some superstable and superregular energy U (or regular with $\sum_k \psi(|k|)<A$). Then $h(\mu,\nu)$ is finite iff $E_\mu|U_i(x)|<\infty$, and the following formula holds*

$$h(\mu,\nu)=e(\mu)-s(\mu)+p=e(\mu)-s(\mu)-(e(\nu)-s(\nu)).$$

Proof. If $s(\mu) > -\infty$, then $\mu|_{\mathscr{B}(V)} \ll d\omega_V$. Therefore

$$|V|^{-1} H_V(\mu, \nu) = |V|^{-1} E_\mu(\log \mu_V(x)) - |V|^{-1} E_\mu(\log \nu_V(x)).$$

Because of Theorem (2.10) the first term converges to $-s(\mu)$. If we replace $\log \nu_V(x)$ in the second term by the right hand side of (4.1), the proof is completed by applying the Theorems (4.2), (2.5) and (3.14) respectively (3.15). □

The following corollary is also in Pirlot [7].

(4.8) **Corollary.** *Let μ be stationary with $E_\mu(X_i^2) < \infty$, and let ν be in $\mathscr{G}_0(U)$ for some superstable and superregular energy U (or regular with $\sum_k \psi(|k|) < A$). Then $e(\mu) - s(\mu) = -p = \inf\{e(\lambda) - s(\lambda), \lambda$ stationary$\}$ iff $h(\mu, \nu) = 0$. In particular $h(\mu, \nu) = 0$ if also $\mu \in \mathscr{G}_0(U)$.*

In order to complete the variational principle we still have to show that any stationary field minimizing the free energy is in $\mathscr{G}_0(U)$. Because of the above corollary this is equivalent to the claim that $\mu \in \mathscr{G}_0(U)$ if $h(\mu, \nu) = 0$ for a $\nu \in \mathscr{G}_0(U)$. For finite state space or *finite range* of the energy (i.e. $I_{V,W} = 0$ if $d(V, W) > K$ where $d(V, W) = \inf\{|i-j|, i \in V, j \in W\}$) this is true very generally (Föllmer [1], Preston [8], Theorem 7.1). Without these assumptions it is very difficult; the next result covers at least a wide class of superstable and superregular energies with infinite range.

(4.9) **Theorem.** *Suppose μ is stationary with $E_\mu(X_i^2) < \infty$ and ν is in $\mathscr{G}_0(U)$ for some superstable and superregular energy U satisfying the following additional assumptions:*

$$K^d \sum_{n > K} \psi(n)\, n^{d-1} \to 0 \quad \textit{for } K \to \infty, \quad \textit{and} \quad \Phi_V = 0 \quad \textit{for all } |V| > 2.$$

Then $h(\mu, \nu) = 0$ implies that also $\mu \in \mathscr{G}_0(U)$.

Proof. Consider the energies U^K with $U_V^K(x) = \sum_{i \in V} \Phi_i(x) + \frac{1}{2} \sum_{|i-j| < K} \Phi_{ij}(x)$. If K is big enough, U^K is also superstable. The corresponding conditional distributions are denoted by $\pi^{K,V}$, and ν^K is in $\mathscr{G}_0(U^K)$. Obviously $|e(\mu) - e^K(\mu)| \leq \frac{1}{2} E_\mu(X_i^2) \sum_{|j| > K} \psi(|j|)$. Moreover Jensen's inequality and Theorem (1.11) show that

$$\begin{aligned} &\log Z_V(0) - \log Z_V^K(0) \\ &\quad \geqq \int (-U_V(x) + U_V^K(x))\, \pi^{K,V}(x \mid 0)\, d\omega_V(x) \geqq -|V|\, C \sum_{|j| > K} \psi(|j|). \end{aligned}$$

This implies $p^K - p \leqq C \sum_{|j| > K} \psi(|j|)$, so because of Theorem (4.7)

$$h(\mu, \nu^K) = h(\mu, \nu) + e^K(\mu) - e(\mu) + p^K - p \leqq \text{const.} \sum_{|j| > K} \psi(|j|),$$

i.e. $K^d h(\mu, \nu^K) \to 0$ for $K \to \infty$.

Now we are going to prove the theorem with the same arguments as Preston [8], p. 115–122, in the case of finite range. We will use the same notations too. Let $A \in \mathscr{B}(V)$, $B \in \mathscr{B}(W)$, $V \cap W = \emptyset$. Then we have (cf. Preston, Lemmas 7.1 and 7.2):

$$\mu(A \cap B) - \int_B \int_A \pi^{K,V}(x \mid y)\, d\omega_V(x)\, d\mu_{V^c}(y) = \int_{A \cap B} (q^K_{V,V'} - 1)\, g^K_{V'}\, d\nu^K$$

for all $V' \supseteq W' = W \cup \{i \notin V, |i-j| \leqq K$ for some $j \in V\}$ (q and g are defined in [8]). We want to show that the right hand side above tends to zero for $K \to \infty$ which will prove the theorem.

Let $\varepsilon > 0$ be given. Choose K_0 such that $h(\mu, \nu^K) < \varepsilon |W'|^{-1}$ for $K > K_0$. As in [8], Lemma 7.6, it is then shown that there is a $V' \supseteq W'$ such that $\int \psi(q^K_{V,V'})\, g^K_{V'}\, d\nu^K < \varepsilon$. Lemma 7.3 of [8] implies then that

$$|\mu(A \cap B) - \int_B \int_A \pi^{K,V}(x \mid y)\, d\omega_V(x)\, d\mu_{V^c}(y)| < \varepsilon' \quad \text{for all } K > K_0. \quad \square$$

(4.10). *Remarks.* i) Superregularity is satisfied if $\psi(k) \sim k^{-d-\varepsilon}$, but the condition of Theorem (4.9) requests a decay of ψ of the order $-(2d+\varepsilon)$.

ii) Preston [9] gives a sufficient conditions under which $h(\mu, \nu) = 0$ with $\nu \in \mathscr{G}_0(U)$ implies that also $\mu \in \mathscr{G}_0(U)$. It is a tightness condition on $\pi^V(x \mid y)$ for different y and all V and it can be applied also in other situations, e.g. point processes, but we were not able to prove that it is satisfied in our situation for all superstable and superregular energies.

Note Added in Proof

After submission of this article an error in the paper of Lebowitz-Presutti [5] was found, see the erratum in Comm. Math. Phys. 78, 1, p. 151 (1980). It appeared that the existence of equilibria cannot be proved with superstability and regularity alone. The additional conditions proposed to remedy this situation are very similar to the ones used in our Theorems (3.14) and (3.15), though our hypotheses are still a little bit stronger.

References

1. Föllmer, H.: On entropy and information gain in random fields. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 207–217 (1973)
2. Föllmer, H.: Phase transition and Martin boundary. In Sém. de Prob. IX, Lecture Notes in Mathematics **465**, 305–318. Berlin-Heidelberg-New York: Springer 1975
3. Künsch, H.: Gaussian Markov random fields. J. Fac. Sci. Univ. Tokyo Sec. IA Math. **26**, 53–73 (1979)
4. Lanford, III O.E., Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. Comm. Math. Phys. **13**, 194–215 (1969)
5. Lebowitz, J.L., Presutti, E.: Statistical mechanics of systems of unbounded spins. Comm. Math. Phys. **50**, 195–218 (1976)
6. Nguyen, X.X., Zessin, H.: Ergodic theorems for spatial processes. Z. Wahrscheinlichkeitstheorie verw. Gebiete **48**, 133–158 (1979)
7. Pirlot, M.: A strong variational principle for continuous spin systems. J. Appl. Probability **17**, 47–58 (1980)

8. Preston, C.: Random fields. Lecture Notes in Mathematics **534**. Berlin-Heidelberg-New York: Springer 1976
9. Preston, C.: A note on conditions under which Lemma 7.2 of Preston, Random fields, can be obtained. Private communication (1979)
10. Rosanov, Yu.A.: On Gaussian fields with given conditional distributions. Theory Probability Appl. **12**, 381–391 (1967)
11. Ruelle, D.: Superstable interactions in classical statistical mechanics. Comm. Math. Phys. **18**, 127–159 (1970)
12. Ruelle, D.: Probability estimates for continuous spin systems. Comm. Math. Phys. **50**, 189–194 (1976)
13. Spitzer, F.: Champs de Markov Gaussiens. Ecole d'été de prob. Saint Flour III, Lecture Notes in Mathematics **390**, 179–185. Berlin-Heidelberg-New York: Springer 1974
14. Tempel'man, A.A.: Ergodic theorems for general dynamical systems. Transactions Moscow Math. Soc. **26**, 94–132 (1972)

Received August 11, 1980