Commun. Math. Phys. 139, 83–101 (1991)

Communications in Mathematical Physics

Microcanonical Distributions for Lattice Gases

Jean-Dominique Deuschel[1], Daniel W. Stroock[2,*], and Hans Zessin[3]

[1] Mathematics, E.T.H. Zentrum, CH-8092 Zurich, Switzerland
[2] M.I.T., rm. 2-272, Cambridge, MA 02140, USA
[3] Mathematik, Universität Bielefeld, W-4900 Bielefeld, FRG

Received October 17, 1990; in revised form November 15, 1990

Abstract. In this article, a large deviation principle (cf. Theorem 1.3) for the empirical distribution functional is applied to prove a rather general version of Boltzmann's principle (cf. Theorem 3.5) for models with shift-invariant, finite range potentials. The final section contains an application of these considerations to the two dimensional Ising model at sub-critical temperature.

1. A Large Deviation Principle for Lattice Systems

In this section we will prove a large deviation theorem for families of random variables indexed by points on a square lattice. (Related earlier results in this direction can be found in [C, F0, and O].) Thus, let $\mathbb{Z}^d$ be the d-dimensional square lattice. We will write $\Lambda \subset\subset \mathbb{Z}^d$ if Λ is a non-empty finite subset of $\mathbb{Z}^d$ and use $|\Lambda| \in \mathbb{Z}^+$ to denote the cardinality of Λ. Also, for $R \in \mathbb{Z}^+$ and $\Lambda \subset\subset \mathbb{Z}^d$, we define

$$\Lambda(R) \equiv \{\mathbf{k} \in \mathbb{Z}^d : |\mathbf{k} - \Lambda| \leqq R\} \quad \text{and} \quad \partial_R \Lambda \equiv \Lambda(R) \backslash \Lambda$$

to be, respectively, the R-hull and R-boundary of Λ. (Throughout, $|\mathbf{k}| \equiv \max_{i \leqq i \leqq d} |k_i|$.)

Next, let E be a Polish space, $\mathscr{B}_E$ the Borel field over E, and $\Omega = E^{\mathbb{Z}^d}$. We give Ω the product topology, and use $\mathscr{B}_\Omega$ to denote the associated Borel field over Ω. Given a non-empty $\Lambda \subseteqq \mathbb{Z}^d$ and $\mathbf{x} \in \Omega, \mathbf{x}_\Lambda$ will denote the element of E^Λ obtained by restricting $\mathbf{x}$ to $\Lambda, \mathscr{B}_\Lambda$ is the σ-algebra over Ω generated by the projection map $\mathbf{x} \in \Omega \to \mathbf{x}_\Lambda \in E^\Lambda$ (of course, $\mathscr{B}_\Omega = \mathscr{B}_{\mathbb{Z}^d}$), $B_\Lambda(\Omega; \mathbb{R})$ is the set of bounded $\mathbb{R}$-valued, $\mathscr{B}_\Lambda$-measurable functions on Ω, and $C_{\Lambda,b}(\Omega; \mathbb{R})$ is the subset of continuous elements of $B_\Lambda(\Omega; \mathbb{R})$. When $\Lambda = \mathbb{Z}^d$, we will simply write $B(\Omega; \mathbb{R})$ for $B_{\mathbb{Z}^d}(\Omega; \mathbb{R})$ and $C_b(\Omega; \mathbb{R})$

* The first two authors acknowledge support from, respectively, the grants NSF DMS-8802667 and NSF DMS-8913328 & DAAL 03-86-K-0171

for $C_{\mathbb{Z}^d,b}(\Omega;\mathbb{R})$; and we set

$$B_0(\Omega;\mathbb{R}) \equiv \bigcup_{\Lambda \subset\subset \mathbb{Z}^d} B_\Lambda(\Omega;\mathbb{R}) \quad \text{and} \quad C_0(\Omega;\mathbb{R}) \equiv \bigcup_{\Lambda \subset\subset \mathbb{Z}^d} C_{\Lambda,b}(\Omega;\mathbb{R}).$$

(That is, $B_0(\Omega;\mathbb{R})$ and $C_0(\Omega;\mathbb{R})$, are the spaces of "local" bounded measurable, respectively bounded continuous, functions on Ω.) Finally, let $\mathbf{M}_1(\Omega)$ be the space of probability measures on $(\Omega, \mathscr{B}_\Omega)$, and set

$$\langle f, \mu \rangle = \int_\Omega f \, d\mu \quad \text{for} \quad f \in B(\Omega;\mathbb{R}) \quad \text{and} \quad \mu \in \mathbf{M}_1(\Omega).$$

For technical reasons, we will have to consider two topologies on $\mathbf{M}_1(\Omega)$: the local strong one generated by the maps $\mu \in \mathbf{M}_1(\Omega) \mapsto \langle f, \mu \rangle \in \mathbb{R}$ as f runs over $B_0(\Omega;\mathbb{R})$, and the weak one which is generated by the same maps when f runs over $C_0(\Omega;\mathbb{R})$ (or, equivalently, over $C_b(\Omega;\mathbb{R})$). Unless it is explicitly stated to the contrary, *topological considerations on* $\mathbf{M}_1(\Omega)$ *will be with respect to the local strong topology.* On the other hand, the measurable structure on $\mathbf{M}_1(\Omega)$ will always be the one determined by the Borel field for the weak topology (which is, of course, the same σ-algebra as the one generated by the same maps as the ones used to generate either the local strong or the weak topologies).

For each $\mathbf{k} \in \mathbb{Z}^d$, $\theta^{\mathbf{k}}: \Omega \to \Omega$ is the **shift transformation** determined by

$$(\theta^{\mathbf{k}}\mathbf{x})_{\mathbf{j}} = x_{\mathbf{k}+\mathbf{j}} \quad \text{for all } \mathbf{j} \in \mathbb{Z}^d \quad \text{and } \mathbf{x} \in \Omega;$$

and we use $\mathbf{M}_1^{\mathrm{S}}(\Omega)$ to denote the subset of $\nu \in \mathbf{M}_1(\Omega)$ which are shift-invariant (i.e., $\nu = \nu \circ (\theta^{\mathbf{k}})^{-1}$, $\mathbf{k} \in \mathbb{Z}^d$.) Clearly, $\mathbf{M}_1^{\mathrm{S}}(\Omega)$ is closed in the weak (and therefore also the local strong) topology.

For any $\emptyset \neq \Lambda \subseteq \mathbb{Z}^d$ and any $\mu \in \mathbf{M}_1(\Omega)$, let μ_Λ denote the marginal distribution of $\mathbf{y} \in \Omega \mapsto \mathbf{y}_\Lambda \in E^\Lambda$ under μ. Next, given a second $\nu \in \mathbf{M}_1(\Omega)$, define the **entropy** $\mathbf{H}_\Lambda(\mu|\nu)$ **of** μ **relative to** ν **on** Λ by

$$\mathbf{H}_\Lambda(\mu|\nu) = \begin{cases} \int_{E^\Lambda} f_\Lambda \log f_\Lambda \, d\nu_\Lambda & \text{if} \quad \nu_\Lambda \ll \mu_\Lambda \quad \text{and} \quad f_\Lambda = \dfrac{d\mu_\Lambda}{d\nu_\Lambda} \\ \infty & \text{otherwise.} \end{cases}$$

Finally, for $n \in \mathbb{Z}^+$, let V_n denote the cube $[-n, n]^d$ and define

$$\bar{\mathbf{h}}(\mu|\nu) \equiv \overline{\lim_{n\to\infty}} \frac{1}{|V_n|} \mathbf{H}_{V_n}(\mu|\nu) \quad \text{and} \quad \underline{\mathbf{h}}(\mu|\nu) \equiv \underline{\lim_{n\to\infty}} \frac{1}{|V_n|} \mathbf{H}_{V_n}(\mu|\nu)$$

when $\nu, \mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$, and take $\bar{\mathbf{h}}(\mu|\nu) = \underline{\mathbf{h}}(\mu|\nu) = \infty$ when either μ or ν is not shift-invariant. If they coincide, we call

$$\mathbf{h}(\mu|\nu) = \underline{\mathbf{h}}(\mu|\nu) = \bar{\mathbf{h}}(\mu|\nu) \tag{1.1}$$

the **specific entropy of** μ **relative to** ν. Note that $\underline{\mathbf{h}}(\cdot|\nu)$ is concave (cf. Exercise 4.4.41 of [DS]) whereas $\bar{\mathbf{h}}(\cdot|\nu)$ is convex; thus the specific entropy $\mathbf{h}(\cdot|\nu)$ is affine on the set where $\bar{\mathbf{h}}(\cdot|\nu)$ and $\underline{\mathbf{h}}(\cdot|\nu)$ are equal.

Given $P \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ and $R \in \mathbb{N}$, we will say that P is R-**mixing** if, for some $M \in [0, \infty)$ and all cubes $Q \subset\subset \mathbb{Z}^d$,

$$\exp[-M|\partial_R Q|] \leq \frac{\int_\Omega f(\mathbf{x}) g(\mathbf{x}) P(d\mathbf{x})}{\int_\Omega f(\mathbf{x}) P(d\mathbf{x}) \int_\Omega g(\mathbf{x}) P(d\mathbf{x})} \leq \exp[M|\partial_R Q|] \tag{1.2}$$

whenever $f\in B_Q(\Omega;(0,\infty))$ and $g\in B_{\mathbb{Z}^d\setminus Q(R)}(\Omega;(0,\infty))$. Our goal in this section is to show that if $P\in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ is R-mixing, then the large deviations of the **empirical distribution functional**

$$\mathbf{x}\in\Omega\mapsto \mathbf{R}_n(\mathbf{x})\equiv\frac{1}{|V_n|}\sum_{\mathbf{k}\in V_n}\delta_{\theta^{\mathbf{k}}\mathbf{x}}\in\mathbf{M}_1(\Omega)$$

are governed by the specific entropy function $\mu\in\mathbf{M}_1(\Omega)\mapsto\mathbf{h}(\mu|P)\in[0,\infty]$. That is, we will prove the following theorem.

1.3 Theorem. *Assume that $P\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ is R-mixing. Then $\bar{\mathbf{h}}(\cdot|P)=\underline{\mathbf{h}}(\cdot|P)$ on the whole of $\mathbf{M}_1(\Omega)$, and the level sets of $\mu\in\mathbf{M}_1(\Omega)\mapsto\mathbf{h}(\mu|P)\in[0,\infty]$ are compact. Moreover for every $A\in\mathscr{B}_{\mathbf{M}_1(\Omega)}$,*

$$-\inf_{\mu\in A^0}\mathbf{h}(\mu|P)\leqq\varliminf_{n\to\infty}\frac{1}{|V_n|}\log[P(\mathbf{R}_n\in A)]$$

$$\leqq\varlimsup_{n\to\infty}\frac{1}{|V_n|}\log[P(\mathbf{R}_n\in A)]\leqq-\inf_{\mu\in\bar{A}}\mathbf{h}(\mu|P). \tag{1.4}$$

The strategy behind our proof of Theorem 1.3 is the same as the one on which we based the proof of Theorem 5.4.27 in [DS]. As in [DS], throughout, P is an R-mixing element of $\mathbf{M}_1^{\mathrm{S}}(\Omega)$, M is the constant in (1.2), and Q denotes a generic, non-empty, finite cube in $\mathbb{Z}^d$.

Step 1. (*Upper Bound for Finite Cubes*) For $f\in B_0(\Omega,\mathbb{R})$, define

$$\mathbf{M}_n(f)=\int_\Omega\exp\left[\sum_{\mathbf{k}\in V_n}f(\theta^{\mathbf{k}}\mathbf{x})\right]P(d\mathbf{x}),\quad n\in\mathbb{Z}^+,$$

and

$$\bar{\Lambda}(f)=\varlimsup_{n\to\infty}\frac{1}{|V_n|}\log\mathbf{M}_n(f).$$

Next, for each Q, define $I_Q:\mathbf{M}_1(\Omega)\mapsto[0,\infty]$ by

$$I_Q(\mu)=\sup\{\langle f,\mu\rangle-\bar{\Lambda}(f):f\in B_Q(\Omega;\mathbb{R})\}.$$

Our goal in this step is to prove that:

$$\mathbf{H}_Q(\mu|P)\leqq|Q(R)|I_Q(\mu)+M|\partial_RQ|,\quad\mu\in\mathbf{M}_1(\Omega), \tag{1.5}$$

and

$$\varlimsup_{n\to\infty}\frac{1}{|V_n|}\log[P((\mathbf{R}_n)_Q\in F_Q)]\leqq-\inf\{I_Q(\mu):\mu_Q\in F_Q\} \tag{1.6}$$

for every closed set F_Q in $\mathbf{M}_1(E^Q)$. Notice, that as a consequence of (1.5), we will know that

$$\{\mu_Q:I_Q(\mu)\leqq L\}\subset\subset\mathbf{M}_1(E^Q)\quad\text{for each}\quad L\in(0,\infty). \tag{1.7}$$

The proofs of (1.5) and (1.6) depend on the following estimate:

$$\mathbf{M}_n(f)\leqq\exp[M\beta(n,Q)|\partial_RQ|\,|V_n|]\left(\int_\Omega e^{|Q(R)|\,|f(\mathbf{x})|}P(d\mathbf{x})\right)\beta(n,Q)|V_n|,$$

where

$$\beta(n,Q)\equiv\left(\frac{1}{n}+\frac{1}{|Q(R)|^{1/d}}\right)^{d} \tag{1.8}$$

for $f\in B_Q(\Omega;\mathbb{R})$ and $n\geqq|Q(R)|^{1/d}$. In proving (1.8), we may and will assume both that $Q=[0,l-1]^d$ and that $f\geqq 0$. Now, take m to be the smallest integer larger than $\frac{n}{l+2R}$, and note that

$$\begin{aligned}\mathbf{M}_n(f)&\leqq\int_\Omega\exp\left[\sum_{\mathbf{j}\in Q(R)}\sum_{\mathbf{k}\in V_n}f(\theta^{\mathbf{j}+(l+R)\mathbf{k}}\mathbf{x})\right]P(d\mathbf{x})\\&\leqq\int_\Omega\exp\left[\sum_{\mathbf{k}\in V_m}|Q(R)|f(\theta^{(l+R)\mathbf{k}}\mathbf{x})\right]P(d\mathbf{x})\\&\leqq\exp[M|V_m|\,|\partial_R Q|]\left(\int_\Omega e^{|Q(R)|f(\mathbf{x})}P(d\mathbf{x})\right)^{|V_m|},\end{aligned}$$

where the passage to the second line is an application of Jensen's inequality plus shift-invariance and the third line comes from the R-mixing condition. Clearly, this proves (1.8).

From (1.8), we see that

$$\langle f,\mu\rangle-\bar{\Lambda}(f)+\frac{M|\partial_R Q|}{|Q(R)|}\geqq\frac{1}{|Q(R)|}\left(|Q(R)|\langle f,\mu\rangle-\log\left(\int_\Omega e^{|Q(R)|f(\mathbf{x})}P(d\mathbf{x})\right)\right),$$

first for non-negative $f\in B_Q(\Omega;\mathbb{R})$ and then for general ones. Hence, by taking the supremum over $f\in B_Q(\Omega;\mathbb{R})$ and using the extremal expression for relative entropy given in Lemma 3.2.13 of [DS], we arrive at (1.5). We will first prove (1.6) with respect to the *weak topology* on $\mathbf{M}_1(E^Q)$. To this end, first note (cf. Theorem 2.2.4 and Exercise 5.1.13 in [DS]) that (1.5) is essentially immediate when F_Q is weakly compact in $\mathbf{M}_1(E^Q)$. Thus (cf. Lemma 2.1.5 in [DS]), the general result for weakly closed F_Q will follow as soon as we show that, for each $L\in(0,\infty)$, there is a $C_{L,Q}\subset\subset\mathbf{M}_1(E^Q)$ with the property that

$$\overline{\lim_{n\to\infty}}\frac{1}{|V_n|}\log[P((\mathbf{R}_n)_Q\notin C_{L,Q})]\leqq -L.$$

But, starting from (1.8), this becomes an easy application of Lemma 3.2.7 in [DS]. Finally, again because of (1.8), the extension to F_Q's which are closed in the local strong topology is accomplished by an application of Theorem 3.2.21 in [DS].

Step 2 (*The General Upper Bound*). Define

$$I(\mu)=\sup_{n\in\mathbb{Z}^+}I_{V_n}(\mu),\quad \mu\in\mathbf{M}_1(\Omega).$$

Given (1.5) and (1.6), it is an easy matter to check (cf. Exercise 2.1.21 and Theorem 3.2.21 in [DS]) that I has compact level sets and that the upper bound in (1.4) holds when $\mathbf{h}(\mu|P)$ is replaced by $I(\mu)$. Thus, to complete the proof of the upper bound in (1.4), we still need to check that

$$I(\mu)=\begin{cases}\infty & \text{if } \mu\in\mathbf{M}_1(\Omega)\setminus\mathbf{M}_1^{\mathrm{S}}(\Omega)\\ \lim\limits_{n\to\infty}\dfrac{1}{|V_n|}\mathbf{H}_{V_n}(\mu|P) & \text{if } \mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\end{cases}. \tag{1.9}$$

The proof of the first part of (1.9) is easy (cf. the verification of (5.4.16) in [DS]). As for the case when $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$, we use (1.8) to see that

$$\overline{\lim_{n\to\infty}}\frac{1}{|V_n|}\mathbf{H}_{V_n}(\mu|P)\leqq I(\mu)$$

for any $\mu\in\mathbf{M}_1(\Omega)$, and the argument leading to (5.4.19) in [DS] to see that

$$I(\mu)\leqq\underline{\lim_{n\to\infty}}\frac{1}{|V_n|}\mathbf{H}_{V_n}(\mu|P)$$

when $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$.

Step 3 (*The Lower Bound*). Define

$$J(\mu)=-\inf\left\{\underline{\lim_{n\to\infty}}\frac{1}{|V_n|}\log[P(\mathbf{R}_n\in G)]:G\ni\mu\right\}$$

as G runs over open sets in $\mathbf{M}_1(\Omega)$. Trivially, the lower bound in (1.4) holds when $\mathbf{h}(\mu|P)$ is replaced by $J(\mu)$. Hence, all that remains is to check that

$$J(\mu)\leqq\mathbf{h}(\mu|P). \tag{1.10}$$

Since, when $\mu\in\mathbf{M}_1(\Omega)\setminus\mathbf{M}_1^{\mathrm{S}}(\Omega)$ there is no problem, we will restrict our attention to $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$. But, by exactly the same argument as was given in Lemma 5.4.21 of [DS], (1.10) holds for ergodic $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$. Moreover, by the Ergodic Decomposition Theorem (cf. Theorem 5.2.16 in [DS] or [D]), general $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ can be expressed as a (continuous) convex combination of ergodic elements of $\mathbf{M}_1^{\mathrm{S}}(\Omega)$. Thus (cf. Lemma 5.4.15 in [DS]), (1.10) for general $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ comes down to checking that $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\mapsto\mathbf{h}(\mu|P)$ is affine and that $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\mapsto J(\mu)$ is lower semi-continuous and convex. The first of these is easy (cf. (5.4.23) in [DS]). As for the second, the lower semi-continuity presents no problem; and so it remains only to show that

$$J\left(\frac{\mu_++\mu_-}{2}\right)\leqq\frac{J(\mu_+)+J(\mu_-)}{2}, \tag{1.11}$$

for $\mu_+,\mu_-\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ with $J(\mu_\pm)<\infty$. To this end, let G be an open neighborhood of $\frac{\mu_++\mu_-}{2}$. We can then choose a cube Q and an $\varepsilon>0$ with the property that

$$\frac{B_Q(\mu_-,2\varepsilon)+B_Q(\mu_+,2\varepsilon)}{2}\subseteqq G,$$

where

$$B_Q(\mu,r)\equiv\{\alpha\in\mathbf{M}_1(\Omega):L_Q(\alpha_Q,\mu_Q)<r\}$$

with L_Q a Lévy metric (cf. (3.2.1) in [DS]) on $\mathbf{M}_1(E^Q)$. Next, we introduce the notation l to stand for the side length of Q and

$$E_{\pm} = \{\mathbf{e} \in \mathbb{Z}^d : e_1 = \pm 1 \text{ and } |e_i| = 1 \text{ for } 2 \leqq i \leqq d\}.$$

Notice that, for sufficiently large $n \in \mathbb{Z}^+$:

$$\bigcap_{\mathbf{e} \in E_{\pm}} \{\mathbf{x} : \mathbf{R}_{[n/2]}(\theta^{(l+R)\mathbf{e}}\mathbf{x}) \in B_Q(\mu_{\pm}, \varepsilon)\} \subseteqq \{\mathbf{x} : \mathbf{R}_n(\mathbf{x}) \in B_Q(\mu_{\pm}, 2\varepsilon)\};$$

and therefore, by R-mixing,

$$P(\mathbf{R}_n \in G) \geqq \exp[-2^d M |\partial_R V_n|](P(\mathbf{R}_{[n/2]} \in B_Q(\mu_-, \varepsilon)))^{2^{d-1}} (P(\mathbf{R}_{[n/2]} \in B_Q(\mu_+, \varepsilon)))^{2^{d-1}}.$$

Hence, from the definition of J, we arrive at

$$-\varliminf_{n \to \infty} \frac{1}{|V_n|} \log[P(\mathbf{R}_n \in G)] \leqq \frac{J(\mu_-) + J(\mu_+)}{2}. \quad \square$$

Remark 1.12.

i) As a direct consequence of Theorem 1.3 and Varadhan's Lemma we have

$$\lim_{n \to \infty} \frac{1}{|V_n|} \log \int_{\Omega} \exp[|V_n| \Phi(\mathbf{R}_n)] dP = \sup\{\Phi(\mu) - \mathbf{h}(\mu|P) : \mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)\}$$

for every $\Phi \in C(\mathbf{M}_1(\Omega); \mathbb{R})$ which is measurable and bounded above. In particular if we choose

$$\Phi(\mu) \equiv \langle f, \mu \rangle, \quad \text{for some} \quad f \in B_0(\Omega; \mathbb{R}),$$

then

$$\bar{\Lambda}(f) = \underline{\Lambda}(f) \equiv \varliminf_{n \to \infty} \frac{1}{|V_n|} \log \mathbf{M}_n(f)$$

and satisfies

$$\bar{\Lambda}(f) = \sup\{\langle f, \mu \rangle - \mathbf{h}(\mu|P) : \mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)\}.$$

Actually, if we restrict our attention to Φ's of the form $\mu \in \mathbf{M}_1(\Omega) \mapsto \langle f, \mu \rangle \in \mathbb{R}$, where $f \in B_0(\Omega; \mathbb{R})$, then we get away without the lower bound in (1.4). To be more precise, suppose that $P \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ is a measure (not necessarily R-mixing) for which we can show that $\underline{\mathbf{h}}(\cdot|P) = \bar{\mathbf{h}}(\cdot|P)$, that $\mathbf{h}(\cdot|P)$ has compact level sets, and that, for every closed $F \in \mathscr{B}_{\mathbf{M}_1(\Omega)}$,

$$\varlimsup_{n \to \infty} \frac{1}{|V_n|} \log[P(\mathbf{R}_n \in F)] \leqq -\inf_{\mu \in F} \mathbf{h}(\mu|P). \tag{1.13}$$

One can then show that, for every $f \in B_0(\Omega; \mathbb{R})$,

$$\bar{\Lambda}(f) \equiv \varlimsup_{n \to \infty} \frac{1}{|V_n|} \log \mathbf{M}_n(f) = \varliminf_{n \to \infty} \frac{1}{|V_n|} \log \mathbf{M}_n(f) \equiv \underline{\Lambda}(f)$$

and that

$$\bar{\Lambda}(f) = \sup\{\langle f, \mu \rangle - \mathbf{h}(\mu|P) : \mu \in \mathbf{M}_1^{\mathrm{E}}(\Omega)\}, \tag{1.14}$$

where $\mathbf{M}_1^{\mathrm{E}}(\Omega)$ denotes the subset of $\mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ which are ergodic. Indeed, the form of Varadhan's Lemma in Lemma 2.1.8 of [DS] together with (1.14) implies

$$\begin{aligned}\bar{\Lambda}(f) &\leqq \sup\{\langle f,\mu\rangle - \mathbf{h}(\mu|P):\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\}\\ &\leqq \sup\{\langle f,\mu\rangle - \mathbf{h}(\mu|P):\mu\in\mathbf{M}_1^{\mathrm{E}}(\Omega)\},\end{aligned}$$

where, in the passage to the second line, we have used the fact that $\mathbf{h}(\cdot|P)$ is affine, the Ergodic Decomposition Theorem, and Lemma 5.4.24 of [DS]. At the same time, by the form of Varadhan's Lemma in Lemma 2.1.7 of [DS],

$$\begin{aligned}\underline{\Lambda}(f) &\geqq \sup\{\langle f,\mu\rangle - J(\mu):\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\}\\ &\geqq \sup\{\langle f,\mu\rangle - \mathbf{h}(\mu|P):\mu\in\mathbf{M}_1^{\mathrm{E}}(\Omega)\},\end{aligned}$$

where, in the last inequality, we have used the fact that (cf. Step 3 in the proof of Theorem 1.3) with no further assumption one always has

$$J(\mu)\leqq\mathbf{h}(\mu|P)\quad\text{for}\quad\mu\in\mathbf{M}_1^{\mathrm{E}}(\Omega).$$

ii) For $n\in\mathbb{Z}^+$ and $\mathbf{x}\in\Omega$, let $\mathbf{x}^{(n)}\in\Omega$ be the V_n-periodic element of Ω which coincides with $\mathbf{x}$ on V_n. That is,

$$\mathbf{x}^{(n)}_{V_n}=\mathbf{x}_{V_n}\quad\text{and}\quad\mathbf{x}^{(n)}_{V_n+(2n+1)\mathbf{k}}=\mathbf{x}^{(n)}_{V_n},\quad\mathbf{k}\in\mathbb{Z}^d.$$

Next define the V_n-periodic empirical distribution measures

$$\mathbf{x}\in\Omega\mapsto\tilde{\mathbf{R}}_n(\mathbf{x})\equiv\frac{1}{|V_n|}\sum_{\mathbf{k}\in V_n}\delta_{\theta^{\mathbf{k}}\mathbf{x}^{(n)}}\in\mathbf{M}_1(\Omega).$$

It is an easy matter to check that $\tilde{\mathbf{R}}(\mathbf{x})$ is shift-invariant for each $n\in\mathbb{Z}^+$ and $\mathbf{x}\in\Omega$. In addition, for each $1\leqq m\leqq n$,

$$\|(\mathbf{R}_n(\mathbf{x}))_{V_m}-(\tilde{\mathbf{R}}_n(\mathbf{x}))_{V_m}\|_{\mathrm{var}}\leqq\frac{|\partial_m V_{n-m}|}{|V_n|};$$

from which it is clear that Theorem 1.3 holds equally well with $\mathbf{R}_n$ replaced by $\tilde{\mathbf{R}}_n$.

iii) Let $M:\mathbb{Z}^+\to[1,\infty)$ and $\alpha:\mathbb{Z}^+\to[0,1]$ be functions for which

$$\lim_{n\to\infty}\frac{M(n)}{|V_n|}=1\quad\text{and}\quad\lim_{n\to\infty}\alpha(n)=0,$$

and suppose that $P\in\mathbf{M}_1$ satisfies the condition

$$\bar{\Lambda}(f)\leqq\frac{1}{M(n)}\log\left(\int_\Omega\exp[M(n)f(\mathbf{x})]P(d\mathbf{x})\right)+\alpha(n)\tag{1.15}$$

for all $n\in\mathbb{Z}^+$ and $f\in B_{V_n}(\Omega;[0,\infty))$. Then, using precisely the same argument as in Steps 1 and 2 of the demonstration of Theorem 1.3, one can prove that the specific entropy $\mathbf{h}(\cdot|P)$ exists, has compact level sets, and provides the upper bound in (1.4). Thus, if one does not require the lower bound in (1.4), one can get away with a far less than R-mixing. In particular, the following *hypermixing* condition will do: $P\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ is said to be **hypermixing** if there exists a non-increasing $\delta:\mathbb{Z}^+\to[1,\infty)$ satisfying $\lim_{n\to\infty}\delta(n)=1$ for which

$$\|f_1\cdots f_n\|_{L^1(P)}\leqq\prod_{m=1}^{n}\|f_m\|_{L^{\delta(R)}(P)},\quad f_m\in B_{Q_m}(\Omega;\mathbb{R}),\quad m=1,..n,$$

whenever $n \geqq 2$ and $Q_1, \ldots, Q_n \subset\subset \mathbb{Z}^d$ are cubes with dist $(Q_l, Q_m) > R$, $m \neq l$. Since one can easely show (cf. Lemma 5.4.13 in [DS]) that for any cube $Q \subset\subset \mathbb{Z}^d$, hypermixing implies

$$\bar{\Lambda}(f) \leqq \frac{1}{|Q(R)|\delta(R)} \log\left(\int_\Omega \exp[|Q(R)|\delta(R) f(\mathbf{x})] P(d\mathbf{x}) \right), \quad f \in B_Q(\Omega; \mathbb{R}),$$

it is clear that hypermixing is more than enough to imply that (1.13) holds. In connection with applications of the sort discussed below, this observation may be useful when dealing systems in which the interaction is unbounded.

2. Gibbs' States and the Variational Principle

In this and the next sections we will be discussing the Gibbs' states on Ω which come from a shift-invariant, finite range potential $\mathscr{U}$ and a reference measure $\lambda \in \mathbf{M}_1(E)$. To be more precise, we will say that $\mathscr{U} = \{U_F : F \subset\subset \mathbb{Z}^d\} \subseteqq B_0(\Omega; \mathbb{R})$ is a **shift-invariant, finite range potential** if:

1) U_F is bounded and $\mathscr{B}_F$ measurable for each $F \subset\subset \mathbb{Z}^d$,
2) $U_{\mathbf{k}+F} = U_F \circ \theta^{\mathbf{k}}$ for each $F \subset\subset \mathbb{Z}^d$ and all $\mathbf{k} \in \mathbb{Z}^d$,
3) there is an $R \in \mathbb{Z}^+$ (the range of $\mathscr{U}$) for which $U_F \equiv 0$ whenever $F \ni \mathbf{0}$ and $F \nsubseteq [-R, R]^d$.

Next, let $\lambda \in \mathbf{M}_1(E)$ be a fixed reference measure, set $\boldsymbol{\lambda} = \lambda^{\mathbb{Z}^d}$, and, for a given shift-invariant, finite range potential $\mathscr{U}$ and $\beta \in \mathbb{R}$, construct the family

$$\{\gamma_{\beta,\Lambda}(\cdot|\mathbf{y}) : \mathbf{y} \in \Omega \text{ and } \Lambda \subset\subset \mathbb{Z}^d\} \subseteqq \mathbf{M}_1(\Omega)$$

by the prescription that, for every $f \in B(\Omega; \mathbb{R})$,

$$\int_\Omega f(\mathbf{x}) \gamma_{\beta,\Lambda}(d\mathbf{x}|\mathbf{y}) = \frac{1}{Z_{\beta,\Lambda}(\mathscr{U}, \mathbf{y})} \int_\Omega f(\mathbf{x}_\Lambda \cdot \mathbf{y}_{\Lambda^\circ}) \exp\left[-\beta \sum_{F \cap \Lambda \neq \varnothing} |F| U_F(\mathbf{x}_\Lambda \cdot \mathbf{y}_{\Lambda^\circ}) \right] \boldsymbol{\lambda}(d\mathbf{x}),$$

where, for $\mathbf{x}, \mathbf{y} \in \Omega$, $\mathbf{x}_\Lambda \cdot \mathbf{y}_{\Lambda^\circ}$ is the element of Ω whose restrictions to Λ and Λ^C coincide with those of $\mathbf{x}$ and $\mathbf{y}$, respectively. (The number $Z_{\beta,\Lambda}(\mathscr{U}, \mathbf{y})$ is determined by the condition that $\gamma_{\beta,\Lambda}(\Omega|\mathbf{y}) = 1$.) It is then an easy matter to check that $\{\gamma_{\beta,\Lambda}(\cdot|\mathbf{y}) : \Lambda \subset\subset \mathbb{Z}^d$ and $\mathbf{y} \in \Omega\}$ is a consistent family of regular conditional probabilities in the sense that, for $\Lambda_1 \subset \Lambda_2$,

$$\int_\Omega f(\mathbf{x}) \gamma_{\beta,\Lambda_2}(d\mathbf{x}|\mathbf{y}) = \int_\Omega \left(\int_\Omega f(\mathbf{x}) \gamma_{\beta,\Lambda_1}(d\mathbf{x}|\xi) \right) \gamma_{\beta,\Lambda_2}(d\xi|\mathbf{y});$$

and we will say that $\gamma \in \mathbf{M}_1(\Omega)$ is a **Gibbs' state with potential $\mathscr{U}$ at inverse temperature β** and will write $\gamma \in \mathfrak{G}_\beta(\mathscr{U})$ if, for each $\Lambda \subset\subset \mathbb{Z}^d$, $\mathbf{y} \mapsto \gamma_{\beta,\Lambda}(\cdot|\mathbf{y})$ is a conditional probability distribution of γ given $\mathscr{B}_{\Lambda^\circ}$ (i.e.,

$$\int_\Omega f(\mathbf{x}) \gamma(d\mathbf{x}) = \int_\Omega \left(\int_\Omega f(\mathbf{x}_\Lambda \cdot \mathbf{y}_{\Lambda^\circ}) \gamma_{\beta,\Lambda}(d\mathbf{x}|\mathbf{y}) \right) \gamma(d\mathbf{y})$$

for every $f \in B(\Omega; \mathbb{R})$). Finally, we set $\mathfrak{G}_\beta^S(\mathscr{U}) \equiv \mathfrak{G}_\beta(\mathscr{U}) \cap \mathbf{M}_1^S(\Omega)$.

Note that $\boldsymbol{\lambda}$ is 0-mixing and that, for any shift-invariant potential $\mathscr{U}$ with range R, all shift-invariant elements of $\mathfrak{G}_\beta(\mathscr{U})$ are R-mixing. In particular, this shows that

R-mixing measures *need not be ergodic.* Moreover, given such a potential $\mathscr{U}$ and a $\gamma\in\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})$, one has that

$$M^{-\beta|\partial_R\Lambda|}\gamma_{\beta,\Lambda}(\cdot|\mathbf{y})\leqq\gamma_\Lambda\times\lambda^{\mathbb{Z}^d\setminus\Lambda}\leqq M^{\beta|\partial_R\Lambda|}\gamma_{\beta,\Lambda}(\cdot|\mathbf{y}),$$

for some $M\in[1,\infty)$ and all $\Lambda\subset\subset\mathbb{Z}^d$, $\beta\in\mathbb{R}$, $\mathbf{y}\in\Omega$. Thus not only does each $\gamma\in\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})$ satisfy the large deviation principle in (1.4), but we also have

$$-\inf_{\mu\in A^0}\mathbf{h}(\mu|\gamma)\leqq\varliminf_{n\to\infty}\frac{1}{|V_n|}\log\left[\inf_{\mathbf{y}\in\Omega}\gamma_{\beta,V_n}(\mathbf{R}_n\in A|\mathbf{y})\right]$$
$$\leqq\varlimsup_{n\to\infty}\frac{1}{|V_n|}\log\left[\sup_{\mathbf{y}\in\Omega}\gamma_{\beta,V_n}(\mathbf{R}_n\in A|\mathbf{y})\right]\leqq-\inf_{\mu\in\bar{A}}\mathbf{h}(\mu|\gamma).$$

In particular, this implies that $\mathbf{h}(\cdot|\gamma)$ is independent of the choice of $\gamma\in\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})$.

When E is compact and $\mathscr{U}\subseteqq C_0(\Omega;\mathbb{R})$, the afore-mentioned consistency of the $\gamma_{\beta,\Lambda}(\cdot|\mathbf{y})$'s guarantees that $\mathfrak{G}_\beta(\mathscr{U})\neq\varnothing$. However, because we have not assumed that E is compact and do not want to restrict ourselves to $\mathscr{U}\subseteqq C_0(\Omega;\mathbb{R})$, even the existence of a Gibbs's state is not entirely obvious; and this is one reason why it will be important for us to have an alternative characterization of Gibbs' states in terms of a variational principle. Namely, for $\Lambda\subset\subset\mathbb{Z}^d$, set

$$\mathscr{U}_\Lambda=\sum_{\varnothing\neq F\subseteqq\Lambda}|F|U_F\quad\text{and}\quad\gamma_{\beta,\Lambda}(d\mathbf{x})=\frac{1}{Z_{\beta,\Lambda}(\mathscr{U})}\exp[-\beta\mathscr{U}_\Lambda(\mathbf{x})]\lambda(d\mathbf{x}),$$

where $Z_{\beta,\Lambda}(\mathscr{U})$ is determined by the condition that $\gamma_{\beta,\Lambda}\in\mathbf{M}_1(\Omega)$. Note that if $R>0$ is the range of $\mathscr{U}$, then there is $M>0$ such that

$$\left|\mathscr{U}_\Lambda-\sum_{\mathbf{k}\in\Lambda}\mathscr{U}_\mathbf{k}\right|\leqq M|\partial_R\Lambda|,\quad\Lambda\subset\subset\mathbb{Z}^d,\tag{2.1}$$

where

$$\mathscr{U}_\mathbf{k}\equiv\sum_{F\ni\mathbf{k}}U_F=\mathscr{U}_0\theta^\mathbf{k},\quad\mathbf{k}\in\mathbb{Z}^d.$$

In particular, this means that

$$\lim_{n\to\infty}\frac{1}{|V_n|}\left|\log(Z_{\beta,V_n}(\mathscr{U}))-\log\left(\int_\Omega\exp[|V_n|\langle\mathscr{U}_0,\mathbf{R}_n(\mathbf{x})\rangle]\lambda(d\mathbf{x})\right)\right|=0;$$

and therefore, since λ (as a 0-mixing element of $\mathbf{M}_1^{\mathbf{S}}(\Omega)$) trivially satisfies the hypotheses of Theorem 1.3, Remark 1.12 leads immediately to the expression

$$\mathbf{p}_\beta(\mathscr{U})\equiv\lim_{n\to\infty}\frac{1}{|V_n|}\log(Z_{\beta,V_n}(\mathscr{U}))=\sup\{-\beta\langle\mathscr{U}_0,\mu\rangle-\mathbf{h}(\mu|\lambda):\mu\in\mathbf{M}_1(\Omega)\}\tag{2.2}$$

for the **pressure** $\mathbf{p}_\beta(\mathscr{U})$; and this, in turn, leads to the following characterization of shift-invariant Gibbs' states (cf. [L] for an earlier account of the same principle).

2.3 Theorem (*The Variational Principle*). *For each* $\beta\in\mathbb{R}$, $\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})\equiv\mathfrak{G}_\beta(\mathscr{U})\cap\mathbf{M}_1^{\mathbf{S}}(\Omega)$ *is a non-empty, convex, compact subset of* $\mathbf{M}_1^{\mathbf{S}}(\Omega)$. *Moreover,*

$$\mu\in\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})\Leftrightarrow\mathbf{p}_\beta(\mathscr{U})=-\beta\langle\mathscr{U}_0,\mu\rangle-\mathbf{h}(\mu|\lambda)$$
$$\Leftrightarrow\beta\langle\mathscr{U}_0,\mu\rangle+\mathbf{h}(\mu|\lambda)\leqq\langle\mathscr{U}_0,\nu\rangle+\mathbf{h}(\nu|\lambda),\quad\nu\in\mathbf{M}_1^{\mathbf{S}}(\Omega).\tag{2.4}$$

Finally, for all $\mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ *and* $\gamma \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$, $\mathbf{h}(\mu|\gamma)$ *exists* (i.e., $\underline{\mathbf{h}}(\mu|\gamma) = \bar{\mathbf{h}}(\mu|\gamma)$) *and*

$$\begin{aligned}\mathbf{h}(\mu|\lambda) &= \mathbf{h}(\mu|\gamma) - \beta\langle \mathscr{U}_0, \mu\rangle - \mathbf{p}_\beta(\mathscr{U}) \\ &= \mathbf{h}(\mu|\gamma) + \mathbf{h}(\gamma|\lambda) + \beta(\langle \mathscr{U}_0, \gamma\rangle - \langle \mathscr{U}_0, \mu\rangle). \qquad (2.5)\end{aligned}$$

Proof. We begin with the observation that

$$\mathbf{H}_{V_n}(\mu|\gamma_{\beta,V_n}) = \mathbf{H}_{V_n}(\mu|\lambda) + \log(Z_{\beta,V_n}(\mathscr{U})) + \beta\langle \mathscr{U}_{V_n}, \mu\rangle,$$

which, in conjunction with (2.1) and (2.2), leads immediately to

$$\lim_{n\to\infty} \frac{\mathbf{H}_{V_n}(\mu|\gamma_{\beta,V_n})}{|V_n|} = \mathbf{h}(\mu|\lambda) + \mathbf{p}_\beta(\mathscr{U}) + \beta\langle \mathscr{U}_0, \mu\rangle, \quad \mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega). \qquad (2.6)$$

Clearly, the second equivalence in (2.4) is just a re-statement of the second part of (2.2). In addition, because

$$\begin{aligned}\gamma \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) &\Rightarrow \mathbf{H}_{V_n}(\gamma|\gamma_{\beta,V_n}) \leqq M|\partial_R V_n| \quad \text{for some } M \in (0,\infty) \text{ and all } n \in \mathbb{Z}^+ \\ &\Rightarrow \lim_{n\to\infty} \frac{\mathbf{H}_{V_n}(\mu|\gamma_{\beta,V_n})}{|V_n|} = 0,\end{aligned}$$

the implication

$$\gamma \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) \Rightarrow \mathbf{p}_\beta(\mathscr{U}) = -\beta\langle \mathscr{U}, \gamma\rangle - \mathbf{h}(\gamma|\lambda)$$

is an immediate consequence of (2.6). Furthermore, because $\mathbf{h}(\cdot|\lambda)$ has compact level sets, it is clear both that

$$\{\mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega) : \mathbf{p}_\beta(\mathscr{U}) = -\beta\langle \mathscr{U}, \mu\rangle - \mathbf{h}(\mu|\lambda)\}$$

is convex and compact and that there is at least one $\mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)$ for which the supremum in (2.2) is achieved. Hence, all that remains is to check the implication

$$\mathbf{p}_\beta(\mathscr{U}) = -\beta\langle \mathscr{U}, \mu\rangle - \mathbf{h}(\mu|\lambda) \Rightarrow \mu \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}). \qquad (2.7)$$

In proving (2.7), we will follow a line of reasoning which we have adapted to the present setting from [P]; and for this purpose, we will need a little preparation. Given a $\mu \in \mathbf{M}_1(\Omega)$ and a $\Lambda \subset\subset \mathbb{Z}^d$, let $\mathbf{y} \in \Omega \mapsto \mu_\Lambda(\cdot|\mathbf{y}) \in \mathbf{M}_1(\Omega)$ be a regular conditional probability distribution of μ given $\mathscr{B}_{\mathbb{Z}^d \setminus \Lambda}$. Obviously,

$$\mu \in \mathfrak{G}_\beta(\mathscr{U}) \Leftrightarrow \int_\Omega \mathbf{H}(\mu_\Lambda(\cdot|\mathbf{y})|\gamma_{\beta,\Lambda}(\cdot|\mathbf{y}))\mu(d\mathbf{y}) = 0 \quad \text{for all } \Lambda \subset\subset \mathbb{Z}^d.$$

Next, use $\Gamma_{\beta,\Lambda}(\mu)$ to denote the element of $\mathbf{M}_1(\Omega)$ determined by

$$\int_\Omega f(\mathbf{x})[\Gamma_{\beta,\Lambda}(\mu)](d\mathbf{x}) = \int_\Omega \left(\int_\Omega f(\mathbf{x})\gamma_{\beta,\Lambda}(d\mathbf{x}|\mathbf{y})\right)\mu(d\mathbf{y}), \quad f \in \mathbf{B}(B;\mathbb{R}).$$

Then, by Lemma 4.4.7 in [DS],

$$\int_\Omega \mathbf{H}(\mu_\Lambda(\cdot|\mathbf{y})|\gamma_{\beta,\Lambda}(\cdot|\mathbf{y}))\mu(d\mathbf{y}) = \mathbf{H}(\mu|\Gamma_{\beta,\Lambda}(\mu)) = \lim_{F\nearrow\mathbb{Z}^d} \mathbf{H}_F(\mu|\Gamma_{\beta,\Lambda}(\mu)).$$

Thus, if we set

$$\Psi_\mu(F,\Lambda) = \mathbf{H}_F(\mu|\Gamma_{\beta,\Lambda}(\mu)) \quad \text{for } \varnothing \neq \Lambda \subseteqq F \subset\subset \mathbb{Z}^d,$$

then we see that

$$\mu\in\mathfrak{G}_\beta(\mathscr{U})\Leftrightarrow \lim_{F\nearrow\mathbb{Z}^d}\Psi_\mu(F,\Lambda)=0 \quad\text{for all}\quad \Lambda\subset\subset\mathbb{Z}^d. \tag{2.8}$$

Finally, note that if $\Lambda\subset\subset\mathbb{Z}^d$ and $\Lambda(R)\subseteqq G\subset\subset\mathbb{Z}^d$, then

$$\begin{aligned}(\gamma_{\beta,G})_\Lambda(\cdot|\cdot)&=\gamma_{\beta,\Lambda}(\cdot|\cdot)\\ &=((\gamma_{\beta,G})_F)_\Lambda(\cdot|\cdot)\quad\text{if}\quad \Lambda(R)\subseteqq F\quad\text{and}\quad F(R)\subseteqq G;\end{aligned}$$

and so, by another application of Lemma 4.4.7 in [DS], we see that

$$\Psi_\mu(F,\Lambda)=\mathbf{H}_F(\mu|\gamma_{\beta,G})-\mathbf{H}_{F\setminus\Lambda}(\mu|\gamma_{\beta,G})\quad\text{if}\quad \Lambda(R)\subseteqq F\quad\text{and}\quad F(R)\subseteqq G. \tag{2.9}$$

Now let $\mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)$ satisfying $\mathbf{p}_\beta(\mathscr{U})=-\beta\langle\mathscr{U}_0,\mu\rangle-\mathbf{h}(\mu|\lambda)$ be given. By (2.8), we will know that $\mu\in\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$ as soon as we show that $\lim_{F\nearrow\mathbb{Z}^d}\Psi_\mu(F,\Lambda)=0$ for each $\Lambda\subset\subset\mathbb{Z}^d$; and, because $F\mapsto\Psi_\mu(F;\Lambda)$ is non-decreasing, this comes down to checking that, for each $\Lambda\subset\subset\mathbb{Z}^d$, $\Psi_\mu(V_n,\Lambda)=0$ for all sufficiently large $n\in\mathbb{Z}^+$. Thus, let $\varnothing\neq\Lambda\subset\subset\mathbb{Z}^d$ and an $n\in\mathbb{Z}^+$ satisfying $\Lambda(R)\subseteqq V_n$ be given. For $m\in\mathbb{Z}^+$, set $N(m)=m(2n+1)+n$, let $\mathbf{k}_1,\dots,\mathbf{k}_{(2m+1)^d}$ be an enumeration of $\{(2n+1)\mathbf{j}:\mathbf{j}\in V_m\}$ with $\mathbf{k}_1=\mathbf{0}$, and set

$$\Lambda_i=\mathbf{k}_i+\Lambda,\quad F_i=\mathbf{k}_i+V_n,\quad\text{and}\quad G_i=F_1\cup\dots\cup F_i\quad\text{for}\quad 1\leqq i\leqq(2m+1)^d.$$

Then

$$\begin{aligned}&\mathbf{H}_{V_{N(m)}}(\mu|\gamma_{\beta,V_{N(m)}(R)})\\ &=\mathbf{H}_{V_n}(\mu|\gamma_{\beta,V_{N(m)}(R)})+\sum_{i=2}^{(2m+1)}(\mathbf{H}_{G_i}(\mu|\gamma_{\beta,V_{N(m)}(R)})-\mathbf{H}_{G_{i-1}}(\mu|\gamma_{\beta,V_{N(m)}(R)})).\end{aligned}$$

At the same time, by (2.9),

$$\mathbf{H}_{V_n}(\mu|\gamma_{\beta,V_{N(m)}(R)})=\mathbf{H}_{V_n\setminus\Lambda}(\mu|\gamma_{\beta,V_{N(m)}(R)})+\Psi_\mu(V_n,\Lambda)\geqq\Psi_\mu(V_n,\Lambda)$$

and

$$\begin{aligned}\mathbf{H}_{G_i}(\mu|\gamma_{\beta,V_{N(m)}(R)})-\mathbf{H}_{G_{i-1}}(\mu|\gamma_{\beta,V_{N(m)}(R)})&\geqq\mathbf{H}_{G_i}(\mu|\gamma_{\beta,V_{N(m)}(R)})-\mathbf{H}_{G_i\setminus\Lambda_i}(\mu|\gamma_{\beta,V_{N(m)}(R)})\\ &=\Psi_\mu(G_i,\Lambda_i)\geqq\Psi_\mu(F_i,\Lambda_i)=\Psi_\mu(V_n,\Lambda),\end{aligned}$$

where, in the last equality, we have used shift-invariance. Hence, after combining these, we have that

$$\Psi_\mu(V_n,\Lambda)\leqq\frac{\mathbf{H}_{V_{N(m)}}(\mu|\gamma_{\beta,V_{N(m)}(R)})}{(2m+1)^d}\quad\text{for all}\quad m\in\mathbb{Z}^d;$$

and therefore, by (2.6), we conclude that $\Psi_\mu(V_n,\Lambda)=0$. □

2.10 Remark. Let $\mathfrak{P}(\Omega)$ stand for the set of all shift invariant, finite range potentials on Ω. Then for each $\beta\neq0$, we have

$$\mathbf{h}(\mu|\lambda)=\sup\{-\beta\langle\mathscr{U}_0,\mu\rangle-\mathbf{p}_\beta(\mathscr{U}):\mathscr{U}\in\mathfrak{P}(\Omega)\},\quad \mu\in\mathbf{M}_1^{\mathrm{S}}(\Omega). \tag{2.11}$$

Indeed, since $\mathbf{h}(\cdot|\lambda)$ is lower semicontinuous and convex (in fact, affine)

$$\mathbf{h}(\mu|\lambda)=\sup\{\langle f,\mu\rangle-\sup\{\langle f,\nu\rangle-\mathbf{h}(\nu|\lambda);\nu\in\mathbf{M}_1^{\mathrm{S}}(\Omega)\}:f\in C_0(\Omega;\mathbb{R})\}.$$

At the same time, for each $f \in C_0(\Omega;\mathbb{R})$, one can construct a potential $\mathscr{U} \in \mathfrak{P}(\Omega)$ with the property that

$$\langle f, \mu\rangle = -\beta\langle \mathscr{U}_0, \mu\rangle, \quad \text{for all} \quad \mu \in \mathbf{M}_1^S(\Omega).$$

Namely, assume that $f \in C_{\Lambda_0}(\Omega;\mathbb{R})$ for some $\Lambda_0 \subset\subset \mathbb{Z}^d$ and simply set

$$U_F = \begin{cases} -\dfrac{1}{\beta}\cdot f \circ \theta^{\mathbf{k}} & \text{if} \quad F = \Lambda_0 + \mathbf{k}, \quad \mathbf{k} \in \mathbb{Z}^d \\ 0 & \text{otherwise} \end{cases}.$$

Hence, (2.11) follows immediately when one combines these two observations.

3. The Equivalence of Ensembles

Gibbs' states turn up in statistical mechanics because they are supposed to be *the equilibrium distribution* of the system under consideration, and the reasoning which underlines this supposition is based on the following picture. Think of E as being the *phase* space of an individual particle and of λ as the *Liouville measure* for the dynamics of each particle when it is *free* (i.e., there are no *forces acting on it*). Next, suppose that we place free particles at the lattice $\mathbb{Z}^d$ and have then interact in such a way that the energy produced by the interaction of the particle at $\mathbf{k}$ with the rest of the system is given by

$$\mathscr{U}_{\mathbf{k}}(\mathbf{x}) \equiv \sum_{F \ni \mathbf{k}} U_F(\mathbf{x}) = \mathscr{U}_0 \circ \theta^{\mathbf{k}}(\mathbf{x})$$

when the position (in Ω) of the particles is $\mathbf{x}$. Finally, consider what happens when we allow our interacting system to achieve equilibrium subject only to the constraint that the *average interaction energy* of the particles be some specified number $\bar{U}$. To be more precise, let $n \in \mathbb{Z}^+$ be given and consider the system of particles at the sights in V_n obtained by imposing periodic boundary conditions. When such a system has achieved equilibrium subject only to the constraint that its average interaction energy be $\bar{U}$, one suspects that its distribution should be the measure $\tilde{\mu}_n$ which one gets by conditioning λ on the event

$$\tilde{A}_n(\bar{U}) \equiv \left\{\mathbf{x} \in \Omega : \frac{1}{|V_n|}\sum_{\mathbf{k} \in V_n} \mathscr{U}_{\mathbf{k}}(\mathbf{x}^{(n)}) = \bar{U}\right\}.$$

In the language of statistical mechanics, $\tilde{\mu}_n$ would be called the **microcanonical distribution** of this system and what **Boltzmann's principle**, equivalently, the **principle of equivalence of ensembles**, predicts is that, as $n \to \infty$, $\tilde{\mu}_n$ tends to some $\gamma \in \mathfrak{G}_\beta^S(\mathscr{U})$, where β (the inverse temperature) is determined by the condition that

$$\int_\Omega \mathscr{U}_0(\mathbf{x})\gamma(d\mathbf{x}) = \bar{U}. \tag{3.1}$$

The purpose of this section is to verify the equivalence of ensembles as an application of the theory of large deviations (cf. Theorem 3.5 below). Besides Lanford's ground-breaking article [L], earlier programs of this sort have been carried out by Dobrushin and Tirozzi in the article [DT] and by Georgii in the book [G]. In [DT] the reasoning is based on the Central Limit Theorem whereas the ideas in [G] derive from de Finetti's theory of symmetric random variables. Moreover, Georgii has recently circulated a preprint in which he obtains closely related results by an information theoretic method which was introduced by Csiszár. Thus, at best, all that is being proposed here is a new strategy for handling this sort of question. In fact, the strategy itself is not entirely new, since it has been used already to handle a closely related situation in [SZ] and was carried out when $d=1$ in [S].

In order not to get involved with problems about the existence of regular conditional probability distributions, we will replace the true microcanonical distribution $\tilde{\mu}_n$ by the **approximate microcanonical distribution** $\tilde{\mu}_{n,\delta}$, $\delta\in(0,1]$, which is the conditional distribution of λ given the event (cf. part ii) in Remark 1.12)

$$\tilde{A}_n(\bar{U},\delta)\equiv\left\{\mathbf{x}\in\Omega:\left|\frac{1}{|V_n|}\sum_{\mathbf{k}\in V_n}\mathscr{U}_{\mathbf{k}}(\mathbf{x}^{(n)})-\bar{U}\right|\leqq\delta\right\}, \tag{3.2}$$

and only at the end will we pass to the limit as $\delta\searrow 0$. Note that

$$\tilde{A}_n(\bar{U},\delta)=\{\mathbf{x}\in\Omega:\tilde{\mathbf{R}}_n(\mathbf{x})\in\mathfrak{M}(\bar{U};\delta)\}, \tag{3.3}$$

where

$$\mathfrak{M}(\bar{U},\delta)\equiv\{\nu\in\mathbf{M}_1(\Omega):|\langle\mathscr{U}_0,\nu\rangle-\bar{U}|\leqq\delta\}.$$

In addition, since

$$\lim_{n\to\infty}\left|\int_{\tilde{A}_n(\bar{U},\delta)}f\,d\lambda-\int_{\tilde{A}_n(\bar{U},\delta)}\langle f,\tilde{\mathbf{R}}_n(\mathbf{x})\rangle\lambda(d\mathbf{x})\right|=0 \tag{3.4}$$

for $f\in B_0(\Omega;\mathbb{R})$, when there is no phase transition (i.e. $\mathfrak{G}^{\mathrm{S}}_\beta(\mathscr{U})$ contains precisely one element), we will have reached our goal once we show that

$$\overline{\lim_{\delta\searrow 0}}\;\overline{\lim_{n\to\infty}}\;\lambda(\tilde{\mathbf{R}}_n\notin G\,|\,\tilde{\mathbf{R}}_n\in\mathfrak{M}(\bar{U};\delta))=0$$

for every open neighborhood $G\in\mathscr{B}_{\mathbf{M}_1(\Omega)}$ of $\mathfrak{G}^{\mathrm{S}}_\beta(\mathscr{U})$.

With these preliminaries, we can now state and prove our result.

3.5 Theorem (Boltzmann's principle). *Set*

$$\mathfrak{M}(\bar{U})=\{\mu\in\mathbf{M}^{\mathrm{S}}_1(\Omega):\langle\mathscr{U}_0,\mu\rangle=\bar{U}\}$$

and assume that

$$m(\bar{U})\equiv\inf\{\mathbf{h}(\mu\,|\,\lambda):\mu\in\mathfrak{M}(\bar{U})\}<\infty. \tag{3.6}$$

(Implicit in (3.6) is the assumption that $\mathfrak{M}(\bar{U})\neq\varnothing$.) Then, for each $\delta\in(0,1)$, there is an $N_\delta\in\mathbb{Z}^+$ such that (cf. *(3.2)*)

$$\lambda(\tilde{A}_n(\bar{U},\delta))\geqq\exp\left[-|V_n|\frac{m(\bar{U})}{1-\delta}\right]\quad\textit{for}\quad n\geqq N_\delta.$$

In fact, for any measurable subset A of $\mathbf{M}_1(\Omega)$,

$$\begin{aligned}-\inf_{\nu\in A^0}\mathbf{I}_{\mathscr{U}}(\nu)&\leqq\underline{\lim_{\delta\searrow 0}}\;\underline{\lim_{n\to\infty}}\;\frac{1}{|V_n|}\log[\lambda(\tilde{\mathbf{R}}_n\in A\,|\,\tilde{A}_n(\bar{U},\delta))]\\&\leqq\overline{\lim_{\delta\searrow 0}}\;\overline{\lim_{n\to\infty}}\;\frac{1}{|V_n|}\log[\lambda(\tilde{\mathbf{R}}_n\in A\,|\,\tilde{A}_n(\bar{U},\delta))]\leqq-\inf_{\nu\in\bar{A}}\mathbf{I}_{\mathscr{U}}(\nu),\end{aligned} \tag{3.7}$$

where

$$\mathbf{I}_{\mathscr{U}}(\nu) \equiv \begin{cases} \mathbf{h}(\nu|\lambda) - m(\bar{U}) & \textit{if} \quad \nu \in \mathbf{M}_1^{\mathrm{S}}(\Omega) \quad \textit{and} \quad \langle \mathscr{U}_0, \nu \rangle = \bar{U} \\ \infty & \textit{otherwise.} \end{cases}$$

In particular, if $G \in \mathscr{B}_{\mathbf{M}_1(\Omega)}$ *is an open neighborhood of the set*

$$\mathfrak{F}(\bar{U}) \equiv \{\mu \in \mathfrak{M}(\bar{U}) : \mathbf{h}(\mu|\lambda) = m(\bar{U})\},$$

then

$$\overline{\lim_{\delta \searrow 0}}\, \overline{\lim_{n \to \infty}} \frac{1}{|V_n|} \log [\lambda(\tilde{\mathbf{R}}_n \notin G | \tilde{A}_n(\bar{U}, \delta))] \leqq - \inf_{\nu \notin G} \mathbf{I}_{\mathscr{U}}(\nu) < 0. \tag{3.8}$$

Finally if $\mathscr{U}$ *is not trivial, there is at most one* $\beta \in \mathbb{R}$ *for which*

$$\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) \cap \mathfrak{M}(\bar{U}) \neq \emptyset; \tag{3.9}$$

and, if such a β *exists, then*

$$\mathfrak{F}(\bar{U}) = \{\gamma \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) : \langle \mathscr{U}_0, \gamma \rangle = \bar{U}\}.$$

In particular, if $\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) \cap \mathfrak{M}(\bar{U})$ *contains precisely one element* γ, *then*

$$\lim_{\delta \searrow 0} \overline{\lim_{n \to \infty}} \left| \int_\Omega f\, d\bar{\mu}_{n,\delta} - \int_\Omega f\, d\gamma \right| = \lim_{\delta \searrow 0} \overline{\lim_{n \to \infty}} \left| \frac{\int_{\tilde{A}_n(\bar{U},\delta)} f\, d\lambda}{\lambda(\tilde{A}_n(\bar{U}, \delta))} - \int_\Omega f\, d\gamma \right| = 0 \tag{3.10}$$

for every $f \in C_{\mathrm{b}}(\Omega; \mathbb{R})$.

Proof. We begin with the relationship between $\mathfrak{F}(\mathscr{U})$ and the sets $\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$, $\beta \in \mathbb{R}$. Thus, suppose that $\gamma \ni \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) \cap \mathfrak{M}(\bar{\mathscr{U}})$ for some $\beta \in \mathbb{R}$. Then, by the second equivalence in (2.4), for any $\mu \in \mathfrak{M}(\bar{\mathscr{U}})$, $\mathbf{h}(\mu|\lambda) \geqq \mathbf{h}(\gamma|\lambda)$ and equality holds if and only if $\mu \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$. Hence, there is at most one such β, and, when it exists, $\mathfrak{F}(\bar{U}) = \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}) \cap \mathfrak{M}(\bar{U})$.

In view of the preceding paragraph, all that remains is to check the validity of (3.7). Actually, what we do here is simply point out that (3.7) is an immediate consequence of the *large deviation principle* (1.4) and the Remark 1.12, ii). Indeed, given (1.4) and that remark, one sees immediately that, for every $\delta \in (0, 1)$ and $A \in \mathscr{B}_{\mathbf{M}_1(\Omega)}$,

$$\begin{aligned} -\inf\{\mathbf{h}(\nu|\lambda) : \nu \in A^0 \cap \mathfrak{M}(\bar{U})\} &\leqq \underline{\lim}_{n \to \infty} \frac{1}{|V_n|} \log [\lambda(\tilde{\mathbf{R}}_n \in A \cap \mathfrak{M}(\bar{U}; \delta))] \\ &\leqq \overline{\lim_{n \to \infty}} \frac{1}{|V_n|} \log [\lambda(\tilde{\mathbf{R}}_n \in A \cap \mathfrak{M}(\bar{U}; \delta))] \\ &\leqq -\inf\{\mathbf{h}(\nu|\lambda) : \nu \in \bar{A} \cap \mathfrak{M}(\bar{U}; \delta)\}; \end{aligned}$$

and clearly (3.7) follows immediately from this, (3.3), and the easily verified fact that

$$\inf\{\mathbf{h}(\nu|\lambda) : \nu \in \bar{A} \cap \mathfrak{M}(\bar{U}; \delta)\} \searrow \inf\{\mathbf{h}(\nu|\lambda) : \nu \in \bar{A} \cap \mathfrak{M}(\bar{U})\}$$

as $\delta \searrow 0$. $\square$

3.11 Remark.

i) One advantage to the line of reasoning which we have taken, is that (3.7) together with Varadhan's lemma (cf. Theorem 2.1.10 in [DS]) leads to the statement

$$\lim_{\delta \searrow 0} \overline{\lim_{n\to\infty}} \left| \frac{1}{|V_n|} \log \left[\frac{\int_{\tilde{A}_n(\bar{U},\delta)} \exp[|V_n|\Phi(\tilde{\mathbf{R}}_n)]d\lambda}{\lambda(\tilde{A}_n(\bar{U},\delta))} \right] - \sup\{\langle \Phi, \mu\rangle - \mathbf{I}_{\mathcal{U}}(\mu) : \mu \in \mathbf{M}_1^{\mathrm{S}}(\Omega)\} \right| = 0$$

for every measurable $\Phi \in C(\mathbf{M}_1(\Omega);\mathbb{R})$ which is bounded above.

ii) Let $A_n(\bar{U},\delta)$ be the set which one obtains by removing the tilde from the $\mathbf{R}_n(\mathbf{x})$ on the right-hand side of (3.3). By exactly the same argument as we just used, one can then prove the statements which result from removing the tildes in (3.7). Thus, so far as the empirical measures are concerned, the result is the same whether one considers *periodic boundary conditions* (those corresponding to the quantities with tildes) or *free boundary conditions* (corresponding to taking the tildes away). On the other hand, because (3.4) fails when the tilde is removed from $\mathbf{R}_n(\mathbf{x})$, we do not know how to prove the analogue of (3.10) when the boundary conditions are free.

The results obtained in Theorem 3.5 do not really require the potential to have finite range and hold for all bounded potentials.

4. Further Comments

Let $\mathcal{U}$ be a shift-invariant, infinite range potential, and assume that $\mathcal{U} \subseteqq C_0(\Omega;\mathbb{R})$. Next, referring to Theorem 3.5, assume that $\mathfrak{F}(\bar{U}) \neq \emptyset$; and, for each $n \in \mathbb{Z}^+$ and $\delta > 0$, let $\tilde{M}_{n,\delta} \in \mathbf{M}_1(\mathbf{M}_1(\Omega))$ denote the distribution of $\mathbf{x} \in \Omega \mapsto \tilde{\mathbf{R}}_n(\mathbf{x}) \in \mathbf{M}_1(\Omega)$ under $\lambda(\cdot|\tilde{A}_n(\bar{U},\delta))$. In this section we will discuss the limit behavior of $\{\tilde{M}_{n,\delta}: n \in \mathbb{Z}^+$ and $\delta > 0\}$ as first $n \to \infty$ and then $\delta \searrow 0$.

Throughout this discussion, we will be considering convergence on $\mathbf{M}_1(\mathbf{M}_1(\Omega))$ with respect to the *weak topology built over the weak topology on* $\mathbf{M}_1(\Omega)$ (i.e., the topology on $\mathbf{M}_1(\mathbf{M}_1(\Omega))$ generated by sets of the form

$$\{N \in \mathbf{M}_1(\mathbf{M}_1(\Omega)) : |\langle \Phi, N\rangle - \langle \Phi, M\rangle| < \alpha\}$$

as M runs over $\mathbf{M}_1(\mathbf{M}_1(\Omega))$, Φ over bounded functions on $\mathbf{M}_1(\Omega)$ which are continuous with respect to the weak topology, and α over $(0,\infty)$.) In particular, we will say that $\tilde{M}$ is *a limit point of* $\{\tilde{M}_{n,\delta}: n \in \mathbb{Z}^+$ and $\delta > 0\}$ *as first* $n \to \infty$ *and then* $\delta \searrow 0$ and will write $\tilde{M} \in \mathbf{L}(\bar{U})$ if there exist $\{\delta(l): l \in \mathbb{Z}^+\} \subseteqq (0,\infty)$ and $\{n(k,l): k, l \in \mathbb{Z}^+\} \subseteqq \mathbb{Z}^+$ such that: $\delta(l) \searrow 0$, $n(k,l) \nearrow \infty$ for each $l \in \mathbb{Z}^+$, and

$$\tilde{M} = \lim_{l\to\infty} \lim_{k\to\infty} \tilde{M}_{n(k,l),\delta(l)}$$

in the sense that

$$\lim_{l\to\infty} \overline{\lim_{k\to\infty}} |\langle \Phi, \tilde{M}_{n(k,l),\delta(l)}\rangle - \langle \Phi, \tilde{M}\rangle| = 0$$

for every bounded $\Phi: \mathbf{M}_1(\Omega) \to \mathbb{R}$ which is continuous with respect to the weak topology; and we will say that $\tilde{M}_{n,\delta}$ *tends to* $\tilde{M}$ and will write

$$\tilde{M} = \lim_{\delta\searrow 0} \lim_{n\to\infty} M_{n,\delta} \quad \text{if} \quad \lim_{\delta\searrow 0} \overline{\lim_{n\to\infty}} |\langle \Phi, \tilde{M}_{n,\delta}\rangle - \langle \Phi, \tilde{M}\rangle| = 0$$

for every bounded $\Phi:\mathbf{M}_1(\Omega)\to\mathbb{R}$ which is continuous with respect to the weak topology.

4.1 Lemma. *For each $\delta>0$, the sequence $\{\tilde{M}_{n,\delta}\}_{n=1}^{\infty}$ is relatively compact. Moreover, if*

$$\mathfrak{F}(\bar{U},\delta)=\{\nu\in\mathfrak{M}(\bar{U},\delta):\mathbf{h}(\nu\,|\,\lambda)\leqq m(\bar{U})\}$$

and $\mathbf{L}(\bar{U},\delta)$ is the set of all subsequential limit points of $\{\tilde{M}_{n,\delta}\}_{n=1}^{\infty}$, then

$$\tilde{M}_\delta(\mathfrak{F}(\bar{U},\delta))=1 \quad \textit{for every} \quad \tilde{M}_\delta\in\mathbf{L}(\bar{U},\delta). \tag{4.2}$$

Hence, $\bigcup_{\delta>0}\mathbf{L}(\bar{U},\delta)$ is relatively compact, and $\tilde{M}\in\mathbf{L}(\bar{U})$ if and only if there are sequences $\{\delta(l)\}_{l=1}^{\infty}\subseteqq(0,\infty)$ and $\{\tilde{M}_{\delta(l)}\}_{l=1}^{\infty}\subseteqq\mathbf{M}_1(\mathbf{M}_1(\Omega))$ such that: $\delta(l)\searrow 0$, $\tilde{M}_{\delta(l)}\in\mathbf{L}(\bar{U},\delta(l))$ for each $l\in\mathbb{Z}^+$, and $\tilde{M}_{\delta(l)}\Rightarrow\tilde{M}$ in $\mathbf{M}_1(\mathbf{M}_1(\Omega))$. In particular, $\mathbf{L}(\bar{U})\neq\emptyset$, $\tilde{M}(\mathfrak{F}(\bar{U}))=1$ for every $\tilde{M}\in\mathbf{L}(\bar{U})$, and

$$\tilde{M}=\lim_{\delta\searrow 0}\lim_{n\to\infty}\tilde{M}_{n,\delta} \quad \textit{if and only if} \quad \mathbf{L}(\bar{U})=\{\tilde{M}\}.$$

Proof. Because the level sets of $\mathbf{h}(\cdot\,|\,\lambda)$ are compact in the weak topology, everything comes down to proving that $\{\tilde{M}_{n,\delta}\}_{n=1}^{\infty}$ is relatively compact for each $\delta>0$ and that (4.2) holds. But, by precisely the same argument as was used to prove (3.8), one can show that for every weakly open neighborhood G of $\mathfrak{F}(\bar{U},\delta)$, $\overline{\lim_{n\to\infty}}\,\tilde{M}_{n,\delta}(G^{\mathrm{C}})=0$. Hence, the required relative compactness becomes an application of the Prokohorov–Varadarajan compactness criterion, and (4.2) follows from the fact that (because $\mathscr{U}\subseteqq C_0(\Omega;\mathbb{R})$) $\mathfrak{F}(\bar{U},\delta)$ is weakly closed. □

As a consequence of the considerations in Lemma 4.1, we see that

$$\lim_{\delta\searrow 0}\lim_{n\to\infty}\tilde{M}_{n,\delta}=\delta_\gamma$$

when $\mathfrak{F}(\bar{U})=\{\gamma\}$. On the other hand, when $\mathfrak{F}(\bar{U})$ contains many elements, the situation is not so clear. Nonetheless, we will close with an example which indicates the sort of phenomena which one might expect in general.

Let E be the two point space $\{-1,1\}$, take $d=2$, and therefore $\Omega=\{-1,1\}^{\mathbb{Z}^2}$. Next, let λ be the standard Bernoulli measure (i.e., $\lambda(\{\pm 1\})=\frac{1}{2}$) and consider the Ising potential $\mathscr{U}$ given by

$$U_F(\mathbf{x})=\begin{cases}\dfrac{|x_{\mathbf{k}}-x_{\mathbf{j}}|}{2} & \text{for} \quad F=\{\mathbf{k},\mathbf{j}\} \quad \text{with} \quad |\mathbf{k}-\mathbf{j}|=1\\ 0 & \text{otherwise.}\end{cases}$$

As Lebowitz showed, Onsager's famous result can be used to see that when $\beta>\beta_c=\log(1+\sqrt{2})$ the associated set $\mathfrak{G}_\beta(\mathscr{U})$ contains more than one element. In fact, Aizenman [A] and Higuchi [H] each showed that $\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})=\mathfrak{G}_\beta(\mathscr{U})$ and that

$$\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})=\left\{\frac{1+t}{2}\gamma_\beta^+ +\frac{1-t}{2}\gamma_\beta^- : t\in[-1,1]\right\}, \tag{4.3}$$

where γ_β^+ is characterized as that element of $\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$ for which

$$m^+(\beta)\equiv\langle x_{\mathbf{0}},\gamma_\beta^+\rangle\geqq\langle x_{\mathbf{0}},\gamma\rangle \quad \text{for all} \quad \gamma\in\mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U}), \tag{4.4}$$

and $\gamma_\beta^- = \gamma_\beta^+ \circ \mathbf{T}^{-1}$, where $\mathbf{T}:\Omega \to \Omega$ is the *spin-flip transformation* given by $\mathbf{T}\mathbf{x} = -\mathbf{x}$. In particular, this means that, for a given β, the number $\bar{U}(\beta) \equiv \langle \mathscr{U}_0, \gamma \rangle$ is independent of the choice of $\gamma \in \mathfrak{G}_\beta^{\mathrm{S}}(\mathscr{U})$. In addition, one can show that $\beta \in \mathbb{R} \mapsto \bar{U}(\beta) \in [0, \infty)$ is a strictly decreasing, continuous function.

4.5 Proposition. *Referring to the preceding paragraph, let $\alpha \in (\beta_c, \infty)$ be given, take $\bar{U} = \bar{U}(\alpha)$, and define $\{\tilde{M}_{n,\delta} : n \in \mathbb{Z}^+$ and $\delta > 0\} \subseteq \mathbf{M}_1(\mathbf{M}_1(\Omega))$ accordingly. Then $\mathfrak{F}(\bar{U}) = \mathfrak{G}_\alpha^{\mathrm{S}}(\mathscr{U})$ and*

$$\lim_{\delta \searrow 0} \lim_{n \to \infty} \tilde{M}_{n,\delta} = \tfrac{1}{2}\delta_{\gamma_\alpha^+} + \tfrac{1}{2}\delta_{\gamma_\alpha^-}. \tag{4.6}$$

Proof. That $\mathfrak{F}(\bar{U}) = \mathfrak{G}_\alpha^{\mathrm{S}}(\mathscr{U})$ is clear from Theorem 3.5 and the remark immediately preceding the statement of this proposition. Thus, in view of Lemma 4.1, what we have to show is that every element $\tilde{M}$ of $\mathbf{L}(\bar{U})$ is the measure on the right-hand side of (4.6). To this end, first note that corresponding to $\tilde{M}$ there is a unique $\rho_{\tilde{M}} \in \mathbf{M}_1([-1, 1])$ for which

$$\tilde{M}(A) = \rho_{\tilde{M}}\left(\left\{ t \in [-1, 1] : \frac{1+t}{2}\gamma_\alpha^+ + \frac{1-t}{2}\gamma_\alpha^- \in A \right\}\right).$$

Moreover, since both $\mathscr{U}$ and λ are $\mathbf{T}$-invariant, it is clear that $\rho_{\tilde{M}}$ must be an even measure on $[-1, 1]$. Hence, our problem comes down to checking that $\rho_{\tilde{M}}((-1, 1)) = 0$; which will certainly follow once we show that, for every $0 \leqq a < m^+(\alpha)$,

$$\lim_{\delta \searrow 0} \lim_{n \to \infty} \lambda(|\bar{S}_n| \leqq a | \tilde{A}_{n,l}(\bar{U}, \delta)) = 0, \tag{4.7}$$

where, for each $n \in \mathbb{Z}^+$, we define

$$\mathbf{x} \in \Omega \to \bar{S}_n(\mathbf{x}) \equiv \frac{1}{|V_n|} \sum_{\mathbf{k} \in V_n} x_{\mathbf{k}} = \langle x_0, \tilde{\mathbf{R}}_n(\mathbf{x}) \rangle.$$

In order to prove (4.7), we first partition $\tilde{A}_n(\bar{U}, \delta)$ into the sets $\tilde{A}_{n,l}(\bar{U}, \delta)$, $-2n \leqq l \leqq 2n + 1$, where

$$\tilde{A}_{n,2n+1}(\bar{U}, \delta) = \tilde{A}_n(\bar{U} + (4n+1)\delta_n, \delta_n),$$
$$\tilde{A}_{n,l}(\bar{U}, \delta) = \tilde{A}(\bar{U} + (2l-1)\delta_n, \delta_n) \setminus \tilde{A}_n(\bar{U} + (2l+1)\delta_n, \delta_n), \quad -2n \leqq l \leqq 2n,$$

and $\delta_n = \dfrac{\delta}{4n+2}$. We then have, by Bayes's Law, that

$$\lambda(|\bar{S}_n| \leqq a | \tilde{A}_n(\bar{U}, \delta)) = \sum_{l=-2n}^{2n+1} \lambda(|\bar{S}_n| \leqq a | \tilde{A}_n^l(\bar{U}, \delta)) \lambda(\tilde{A}_{n,l}(\bar{U}, \delta) | \tilde{A}_n(\bar{U}, \delta)),$$

and so (4.7) will follow from

$$\lim_{\delta \searrow 0} \overline{\lim_{n \to \infty}} \max_{-2n \leqq l \leqq 2n+1} \lambda(|\bar{S}_n| \leqq a | \tilde{A}_{n,l}(\bar{U}, \delta)) = 0. \tag{4.8}$$

In proving (4.8), we will make use of the probability measures $\tilde{\gamma}_{\beta, V_n} \in \mathbf{M}_1(\Omega)$ given by:

$$\tilde{\gamma}_{\beta, V_n}(d\mathbf{x}) \equiv \frac{\exp[-\beta |V_n| \langle \mathscr{U}_0, \tilde{\mathbf{R}}_n(\mathbf{x}) \rangle]}{\tilde{Z}_{\beta, V_n}(\mathscr{U})} \lambda(d\mathbf{x}),$$

where $\tilde{Z}_{\beta,V_n}(\mathscr{U})$ is the normalizing constant making $\tilde{\gamma}_{\beta,V_n}$ a probability measure. Clearly,

$$\lambda(\cdot|\tilde{A}_{n,l}(\bar{U},\delta)) \leqq \exp[(2n+1)8\beta\delta]\tilde{\gamma}_{\beta,V_n}(\cdot|\tilde{A}_{n,l}(\bar{U},\delta)) \tag{4.9}$$

for all $\beta\in\mathbb{R}$, $n\in\mathbb{Z}^+$, and $-2n \leqq l \leqq 2n+1$. The importance of (4.9) to us derives from two estimates on the measures $\tilde{\gamma}_{\beta,V_n}$. The first of these is the estimate in [CCS] from which one can show that there exists an $\varepsilon >$ with the properties that

$$m^+(\beta) > a \quad \text{and} \quad \tilde{\gamma}_{\beta,V_n}(|\bar{S}_n| \leqq a) \leqq \frac{e^{-(2n+1)\varepsilon}}{\varepsilon} \quad \text{for all} \quad \beta\in[\alpha-\varepsilon, \alpha+\varepsilon]; \tag{4.10}$$

and the second is the estimate in [N] saying that

$$\varliminf_{n\to\infty} \inf_{|\beta-\alpha|\leqq\varepsilon} \tilde{\gamma}_{\beta,V_n}(|\langle\mathscr{U}_0, \tilde{\mathbf{R}}_n\rangle - \bar{U}(\beta)| < \delta_n) > 0. \tag{4.11}$$

(Actually, the result, Theorem 2 in [CCS], on which (4.10) is based is stated when the boundary conditions are free, not periodic. However, Schonmann assures us that the same techniques apply to the periodic case as well. Also, (4.10) is not explicitly stated in [N], but is implicit in the Central Limit Theorem which Newman derives from Theorem 3 and Proposition 4 of [N].) In particular, since $\beta \mapsto \bar{U}(\beta)$ has a continuous inverse and $\bar{U}(\alpha) = \bar{U}$, we can find a $\delta_0 > 0$ with the property that, for each $\delta\in(0, \delta_0)$, $n\in\mathbb{Z}^+$, and $-2n \leqq l \leqq 2n+1$, there is a unique $\beta(n,l,\delta)\in[\alpha-\varepsilon, \alpha+\varepsilon]$ such that $\bar{U}(\beta(n,l,\delta)) = \bar{U} + (2l-1)\delta_n$ and, therefore,

$$\tilde{A}_{n,l}(\bar{U},\delta) \supseteqq \{\mathbf{x}: |\langle\mathscr{U}_0, \tilde{\mathbf{R}}_n(\mathbf{x})\rangle - \bar{U}(\beta(n,l,\delta))| < \delta_n\}.$$

Hence, by combining (4.9), (4.10), and (4.11), we see that there exists a $K\in(0,\infty)$ such that

$$\begin{aligned}
&\lambda(|\bar{S}_n| \leqq a | \tilde{A}_{n,l}(\bar{U},\delta)) \\
&\quad \leqq \exp[(2n+1)8\beta(n,l,\delta)\delta] \frac{\tilde{\gamma}_{\beta(n,l,\delta),V_n}(|\bar{S}_n| \leqq a)}{\tilde{\gamma}_{\beta(n,l,\delta)}(|\langle\mathscr{U}_0, \tilde{\mathbf{R}}_n\rangle - \bar{U}(\beta(n,l,\delta))| < \delta_n)} \\
&\quad \leqq K \exp[-(2n+1)(\varepsilon - 8(\alpha+\varepsilon)\delta)]
\end{aligned}$$

for all $n\in\mathbb{Z}^+$, $-2n \leqq l \leqq 2n+1$, and $0 < \delta < \delta_0$; and clearly (4.8) follows from this. □

4.12 Remark.

i) It is hardly necessary to point out that (4.6) certainly implies that, as first $n \to \infty$ and then $\delta \searrow 0$, (cf. the notation in Sect. 3) $\tilde{\mu}_{n,\delta}$ tends to $\frac{1}{2}\gamma_\beta^+ + \frac{1}{2}\gamma_\beta^-$. Of course this same conclusion can be reached much more easily and directly by simply observing that every limit of $\tilde{\mu}_{n,\delta}$ must be the (unique) $\gamma\in\mathfrak{G}_\beta^{\mathbf{S}}(\mathscr{U})$ for which $\langle x_0, \gamma\rangle = 0$. In this connection, note that a similar line of reasoning, based on the classification of Gibbs' states in [FP], can be applied to the three dimensional Heisenberg to see that in that case also $\tilde{\mu}_{n,\delta}$ converges, this time to the unique rotation invariant Gibbs' state. On the other hand, we are unable to say what happens to the $\tilde{M}_{n,\delta}$'s in this case.

ii) The result contained in Proposition 4.5 suggests that it is reasonable to expect that elements of $\mathbf{L}(\bar{U})$ ought to be concentrated on the set of extreme points in $\mathfrak{F}(\bar{U})$. Certainly, for those potentials when this is known to be the case, considerations of the sort in Proposition 4.5 would be far simpler. Indeed, if we had known this ahead of time for the two dimensional Ising model, then the argument in the proof of Proposition 4.5 would have ended after we had remarked that (cf. the beginning of that proof) $\rho_{\tilde{M}}$ must be even.

References

[A] Aizenman, M.: Translation invariance and instability of phase coexistence in the two-dimensional Ising system. Commun. Math. Phys. **73**, 83–94 (1980)

[CCS] Chayes, J., Chayes, L., Schonmann, R.: Exponential decay of connectivities in the two-dimensional Ising model. J. Stat. Phys. **49**, 433–445 (1987)

[C] Comets, F.: Grandes déviations pour des champs de Gibbs sur $\mathbb{Z}^d$. C.R. Acad. Sci. Paris, Série I **303**, 511 (1986)

[DS] Deuschel, J.-D., Stroock, D.: Large deviations. Pure and Appl. Math. Series #137. Boston: Academic Press 1989

[DT] Dobrushin, R., Tirozzi, B.: The central limit theorem and the problem of equivalence of ensembles. Commun. Math. Phys. **54**, 173–192 (1977)

[D] Dynkin, E.: Sufficient statistics and extreme points. Ann. Probab. **6**, 705–730 (1978)

[FO] Föllmer, H., Orey, S.: Large deviations for the empirical field of a Gibbs measure. Ann. Prob. **16**, 961–977 (1988)

[FP] Fröhlich J., Pfister, C.-E.: Spin-waves, vortices and the structure of equilibrium states in the classical XY model. Commun. Math. Phys. **89**, 303–327 (1983)

[G] Georgii, H.: Canonical Gibbs Measures, vol. **760**. Lecture Notes Series in Mathematics. Berlin, Heidelberg, New York: Springer 1979

[H] Higuchi, Y.: On the absence of non-translationally invariant Gibbs states for the two-dimensional Ising system. Random Fields: Rigorous Results in Statistical Mechanics and Quantum Field Theory. Fritz, J. J., Lebowitz, J. L., Szàsz, D. (eds). Colloquia Mathematica Societatis Janos Bolyai vol. **27**, pp. 517–534. Amsterdam: North Holland

[L] Lanford, O.: Entropy and equilibrium states in classical statistical mechanics. Statistical Mechanics and Mathematical Physics vol. **20**, pp. 1–133. Lecture Notes in Physics. Berlin, Heidelberg, New York: Springer 1973

[N] Newman, C.: A general central limit theorem for FKG systems. Commun. Math. Phys. **91**, 75–80 (1983)

[O] Olla, S.: Large deviations for Gibbs random fields. Probab. Theory Relat. Fields **77**, 343–357 (1988)

[P] Preston, C.: Random Fields. Lecture Notes in Mathematics vol. **534**. Berlin, Heidelberg, New York: Springer 1976

[S] Stroock, D.: Microcanonical distributions for one dimensional lattice gases, to appear in Proc. of 1989 Lisbon Conf. on Stoch. Anal. Basel: Birkhäuser

[SZ] Stroock, D., Zeitouni, O.: Microcanonical distributions, Gibbs' states, and the equivalence of ensembles. To appear in the Festschrift volume for F. Spitzer

Communicated by A. Jaffe