Probab. Th. Rel. Fields, 77, 343–357 (1988)

Probability Theory and Related Fields

Large Deviations for Gibbs Random Fields*

Stefano Olla**

Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Abstract. A large deviation principle for Gibbs random fields on $\mathbf{Z}^d$ is proven and a corresponding large deviations proof of the Gibbs variational formula is given. A generalization of the Lanford theory of large deviations is also obtained.

0. Introduction

The object of this paper is the study of the large deviation properties of Gibbs measures with respect to the thermodynamic limit. In particular it contains an extension of the Donsker-Varadhan large deviations theory (cf. [5, 6, 15]) to Gibbsian random fields on the lattice $\mathbb{Z}^d$.

The main result is a large deviation principle for the distribution of the empirical measure of a Gibbs field (cf. Definition 1.2 and Theorem 5.3). Large deviation properties for the distributions of the observables of the field can be obtained from this result via the contraction principle (cf. [15]). Thus this work can be seen as a generalization of Lanford's approach to equilibrium statistical mechanics (cf. [9]). The rate function involved is the relative entropy defined by (5.18). It depends only on the interaction and not on the particular Gibbs measure.

It follows that for a given interaction, all the corresponding Gibbs measures have the same large deviation properties (at least at this order of exponential decay of probabilities). Furthermore, for a given interaction, the relative entropy of a Gibbs measure with respect to another Gibbs measure is zero. This will imply that if there exists more than one translation invariant Gibbs measure, then there exist deviations from the thermodynamic limit whose probabilities go to zero slower than exponentially in the volume of the system. Conversely the existence of an observable with a rate function not strictly convex at its

* This work was partially supported by NSF-DMR81-14726

** *Present address:* Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA

minimum implies the existence of more than one translation invariant Gibbs measure (i.e., a phase transition, cf. remark at the end of Sect. 6).

In Sect. 2 we study the properties of the relative entropy of a translation invariant measure with respect to a homogeneous product measure. A quasi-local characterization of the relative entropy (cf. (2.9), used also in [1] and [11]) is particularly useful in the proof of the upper large deviation estimate in Sect. 3. The lower bound for the product measure is proven in Sect. 4. Section 5 contains the large deviation principle for Gibbs measures (Theorem 5.3) and a large deviation proof of the Gibbs variational formula (Theorem 5.1). All these results depend only on the local structure of the Gibbs measure, so they are valid also for non translation invariant Gibbs measures.

The large deviation principles for the distributions of the observables and the connection with Lanford's theory are obtained in Sect. 6 using the contraction principle (cf. [15]).

The results concerning product measures (Sect. 2.4) are established for a complete metrizable state space, while compactness is assumed for the results concerning Gibbs measures (Sects. 5 and 6). In non-Gibbsian situations the strict convexity of the rate function is a typical condition in order to obtain large deviation results (cf. [2, 8, 10]). These kind of conditions are not necessary in the Gibbs case, which permits us to study large deviations also in the presence of phase transition.

For an introduction to the relations between statistical mechanics and large deviations see the interesting book of Ellis [8]. A large deviations proof of the Gibbs variational formula for 1-dimensional ferromagnetic models is also contained in that book.

After the first submission of this work we received the papers of F. Comets [16] in which similar results to those contained in Sects. 3, 4 and 5 are announced. See also H. Föllmer and S. Orey [17] for a different approach using the translation invariance of the Gibbs measure.

1. Notations

We consider the configuration space $\Omega = X^{\mathbf{Z}^d}$ (i.e., $\Omega = \{\omega : \mathbf{Z}^d \to X\}$) where X is a polish space and $d \in \mathbf{N}$. On Ω we consider the product topology.

Given a region $\Lambda \subset \mathbf{Z}^d$ let F_Λ be the σ-algebra on Ω generated by the projections $\omega \in \Omega \to \omega(j) \in X$ for $j \in \Lambda$. We will use the notation $\Omega_\Lambda = \{\omega : \Lambda \to X\} \simeq (\Omega, F_\Lambda)$.

Let $<$ be the natural lexicographical order on $\mathbf{Z}^d$.

Denote by $\mathbf{Z}^d_< = \{x \in \mathbf{Z}^d, x < 0\}$, $F_< = F_{\mathbf{Z}^d_<}$, $F_0 = F_{\{0\}}$.

For $j \in \mathbf{Z}^d$ let θ_j be the translation operator on Ω defined by $(\theta_j \omega)(i) = \omega(i+j)$.

Let $M(X)$ be the space of the probability measures on X and $M_\theta(\Omega)$ be the space of the translation invariant probability measures on Ω, both considered with the weak topology.

Define $\mathbf{Z}^d_+ = \{z \in \mathbf{Z}^d; z_i \geqq 0; i = 1, \ldots, d\}$. For $a \in \mathbf{Z}^d_+$ define the hypercube $\Lambda(a) = \{z \in \mathbf{Z}^d_+; 0 \leqq z_i \leqq a_i\} \subset \mathbf{Z}^d_+$. With $a \to \infty$ we intend $a_i \to \infty$ for any $i = 1, \ldots, d$.

For any $\omega \in \Omega$ define a $\Lambda(a)$-periodized configuration $\omega^a \in \Omega$ in the following way

$$\begin{aligned} \omega^a(j) &= \omega(j) \quad &\text{if } j \in \Lambda(a) \\ \omega^a(j + a_k \mathbf{e}^k) &= \omega^a(j) \quad &\text{for all } j \in \mathbf{Z}^d, \end{aligned} \tag{1.1}$$

where $\mathbf{e}_h^k = \delta_{k,h}$.

Given $\omega \in \Omega$ and $a \in \mathbf{Z}_+^d$, define the empirical distribution of ω in the region $\Lambda(a)$ as

$$R_{\Lambda(a),\omega} = |\Lambda(a)|^{-1} \sum_{j \in \Lambda(a)} \delta_{\theta_j \omega^a}. \tag{1.2}$$

It is easy to check that $R_{\Lambda(a),\omega} \in M_\theta(\Omega)$ and that it depends only on $\{\omega(j), j \in \Lambda(a)\}$.

Let P be a probability measure on Ω, then define the family of probability measures $\{\hat{P}_a, a \in \mathbf{Z}_+^d\}$ on $M_\theta(\Omega)$ as

$$\hat{P}_a(A) = P(R_{\Lambda(a),\omega} \in A).$$

We say that P satisfies a large deviations principle for the empirical distribution if there exists a rate function

$$H:\ M_\theta(\Omega) \to [0, +\infty]$$

such that:

a) H is lower semicontinuous

b) for any closed set $C \subset M_\theta(\Omega)$

$$\lim_{a \to \infty} \sup |\Lambda(a)|^{-1} \log \tilde{P}_a(C) \leqq -\inf_{Q \in C} H(Q)$$

c) for any open set $G \subset M_\theta(\Omega)$

$$\lim_{a \to \infty} \inf |\Lambda(a)|^{-1} \cdot \log \tilde{P}_a(G) \geqq -\inf_{Q \in G} H(Q).$$

In Sects. 3 and 4 we prove a large deviations principle for P a homogeneous product measure, and in Sect. 5 for P a Gibbs measure for a given interaction.

2. Entropy

Given two measures $\nu, \mu \in M(X)$, the relative entropy of ν with respect to μ is defined as

$$h(\nu|\mu) = \sup_{\phi \in B(X)} \{\nu(\phi) - \log \mu(e^\phi)\}. \tag{2.1}$$

The relative entropy $h(\nu|\mu)$ is finite if and only if ν is absolutely continuous with respect to μ and $\log \frac{d\nu}{d\mu} \in L^1(\nu)$. In this situation we have (cf. [5])

$$h(\nu|\mu) = \nu\left(\log \frac{d\nu}{d\mu}\right). \tag{2.2}$$

Given $\mu \in M(X)$ and $Q \in M_\theta(\Omega)$ define

$$H_\mu(Q) = Q(h(q_<(d\omega(0)) \mid \mu(d\omega(0)))) \tag{2.3}$$

where $q_<(d\omega(0))$ is the Q-conditional distribution of $\omega(0)$ given $F_<$.

By the same argument used in [5], Lemma 3.4, there exists a measurable function not depending on Q

$$h\colon (\Omega, F_< \vee F_0) \to \mathbf{R}_+ \qquad \text{such that } H_\mu(Q) = Q(h). \tag{2.4}$$

The proof of (2.4) is based on the same argument used in [5], Lemma 3.4.

By (2.2) if $H_\mu(Q) < \infty$ we have

$$H_\mu(Q) = Q\left(\int_X q_<(d\omega(0)) \log\left(\frac{d q_<}{d\mu}\right)(\omega(0))\right). \tag{2.5}$$

As in (cf. [5], Theorem 3.2) it is possible to give other variational characterizations of $H_\mu(Q)$ as:

$$H_\mu(Q) = \sup_{\phi \in B(F_< \vee F_0)} Q(\phi - \log(\delta_< \otimes_0 \mu)(e^\phi)) \tag{2.6}$$

where the measure $\delta_< \otimes_0 \mu$ is defined by

$$(\delta_< \otimes_0 \mu)\Big(\prod_{j \in Z^d_{\leqq 0}} \psi_j(\omega(j))\Big) = \prod_{j \in Z^d_<} \psi_j(\omega(j))\, \mu(\psi_0(\omega(0)).$$

Let $\{\Lambda_k\}$ be an increasing sequence of finite subsets of $\mathbf{Z}^d_<$ such that $\Lambda_k \uparrow \mathbf{Z}^d_<$ and define

$$Y_{\Lambda_k} = \{\phi \in C(\Omega) \cap B(F_{\Lambda_k} \vee F_0);\ (\delta_< \otimes_{0\,\mu})(e^\phi) \leqq 1\} \tag{2.7}$$

$$Y = U_k Y_{\Lambda_k} \tag{2.8}$$

then using the same argument of ([11], Theorem 2.4) one shows that

$$H_\mu(Q) = \sup_{\phi \in Y} Q(\phi). \tag{2.9}$$

for any choice of the sequence $\{\Lambda_k\}$.

It follows immediately from (2.9) that H_μ is a non negative lower semicontinuous function on $M_\theta(\Omega)$.

This last characterization will be used in the next section to prove the upper estimate for the large deviations.

3. Upper Estimates

This and the following sections concern only the product measure case. Let μ be any probability measure on X and denote $P_\mu = \mu^{Z^d}$. Define the measure $\delta_< \otimes_0 P_\mu$ on Ω as

$$(\delta_< \otimes_0 P_\mu)\Big(\prod_{j \in Z^d} \psi_j(\omega(j))\Big) = \prod_{j \in Z^d_<} \psi_j(\omega(j)) P_\mu\Big(\prod_{j' \in Z^d_{\geqq 0}} \psi_{j'}(\omega(j')).$$

Lemma 3.1. *For any $\phi \in Y$, and for any $a \in \mathbf{Z}^d_+$*

$$(\delta_< \otimes_0 P_\mu)(\exp \sum_{j \in \Lambda(a)} \phi(\theta_j \omega)) \leqq 1. \tag{3.1}$$

Proof. Observe that $\theta_{-a}(\lambda(a) \backslash \{a\}) \subset \mathbf{Z}^d_<$, then

$$\begin{aligned}
&(\delta_< \otimes_0 P_\mu)[(\exp \sum_{j \in \Lambda(a)} \phi(\theta_j \omega))] \\
&= [(\delta_< \otimes_0 P_\mu)(\exp \sum_{j \in \Lambda(a) \backslash \{a\}} \phi(\theta_j \omega))] [(\delta_< \otimes_0 \mu)(\exp \phi(\omega)) \\
&\leqq (\delta_< \otimes_0 P_\mu)(\exp \sum_{j \in \Lambda(a) \backslash \{a\}} \phi(\theta_j \omega))
\end{aligned}$$

and (3.1) follows by iteration on the all set $\Lambda(a)$ following the order given by $<$. □

Lemma 3.2. *Let Γ be a finite subset of $\mathbf{Z}^d_<$, then for any $\phi \in Y_\Gamma$*

$$P_\mu(\exp\{|\Lambda(a)| R_{\Lambda(a),\omega}(\phi)\}) \leqq \exp(2 \|\phi\|_\infty |V_\Gamma(a)|) \tag{3.2}$$

where $V_\Gamma(a) = \{j \in \Lambda(a);\ \Gamma + j \not\subset \Lambda(a)\}$.

Proof. The localization of ϕ on Γ implies that for any $\omega \in \Omega$:

$$|\sum_{j \in \Lambda(a)} [\phi(\theta_j \omega) - \phi(\theta_j \omega^a)]| \leqq 2 \|\phi\|_\infty |V_\Gamma(a)|, \tag{3.3}$$

then (3.2) follows from Lemma 3.1. □

Lemma 3.3. *For any set $A \subset M_\theta(\Omega)$ and any finite subset $\Gamma \in \mathbf{Z}^d_<$:*

$$\limsup_{a \to \infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A) \leqq -\sup_{\phi \in Y_\Gamma} \inf_{Q \in A} Q(\phi). \tag{3.4}$$

Proof. By (3.2) for any $\phi \in Y_\Gamma$ (defined by (2.7)) and by Chebyshev inequality

$$P_\mu(R_{\Lambda(a),\omega} \in A) \leqq \exp(-|\Lambda(a)| \inf_{Q \in A} Q(\phi)) \exp(2 \|\phi\|_\infty |V_\Gamma(a)|)$$

then taking the logarithm of both parts and dividing by the volume

$$|\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A) \leqq -\inf_{Q \in A} Q(\phi) + 2 \|\phi\|_\infty \frac{|V_\Gamma(a)|}{|\Lambda(a)|}.$$

By the finiteness of Γ we have that

$$\lim_{a \to \infty} \frac{|V_\Gamma(a)|}{|\Lambda(a)|} = 0.$$

Then after the limit as $a \to \infty$ it is possible to take the sup over all ϕ in Y_Γ and (3.4) follows. □

Lemma 3.4. *For any compact set $A \subset M_\theta(\Omega)$*

$$\limsup_{a\to\infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A) \leqq -\inf_{Q\in A} H_\mu(Q). \tag{3.5}$$

Proof. Consider an increasing sequence $\{\Gamma_k\}$ of subset of $\mathbf{Z}^d_<$ such that $\Gamma_k \uparrow \mathbf{Z}^d$. Then (3.5) follows by the lower semicontinuity of $H_\mu(Q)$, the argument used in ([5 (IV)] Theorem 4.2 and 4.3) and the variational formula (2.9). □

Theorem 3.5. *For any closed set $A \subset M_\theta(\Omega)$*

$$\limsup_{a\to\infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A) \leqq -\inf_{Q\in A} H_\mu(Q). \tag{3.6}$$

Proof. For any $a \in \mathbf{Z}^d_+$ and any $\omega \in \Omega$ let

$$T_{\Lambda(a),\omega} = |\Lambda(a)|^{-1} \sum_{j\in\Lambda(a)} \delta_{\omega(j)} \in M(X) \tag{3.7}$$

which is the marginal of $R_{\Lambda(a),\omega}$.

By Lemma 3.32 of [13], for any $L>0$ there exists a compact set $C_L \subset M(X)$ such that

$$\limsup_{a\to\infty} |\Lambda(a)|^{-1} \log P_\mu(T_{\Lambda(a),\omega} \in C_L^c) \leqq -L. \tag{3.8}$$

Let $\tilde{C}_L = \{Q \in M_\theta(\Omega); Q_0 \in C_L\}$, where Q_0 is the marginal of Q, then $\tilde{C}_L$ is tight for any $L>0$.

So we have

$$P_\mu(R_{\Lambda(a),\omega} \in A) \leqq P_\mu(R_{\Lambda(a),\omega} \in A \cap \tilde{C}_L) + P_\mu(T_{\Lambda(a),\omega} \in C_L^c).$$

Now $A \cap \tilde{C}_L$ is compact, then by (3.5) and (3.8)

$$\limsup_{a\to\infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A)$$
$$\leqq \max\{-\inf_{A\cap\tilde{C}_L} H_\mu(q), -L\} \leqq \max\{-\inf_A H_\mu(Q), -L\}$$

and the result follows taking L large enough. □

4. Lower Estimates

Lemma 4.1. *Let $Q \in M_\theta(\Omega)$ be ergodic and such that $H_\mu(Q) < +\infty$. Then for any open neighborhood N of $Q \in M_\theta(\Omega)$ we have*

$$\liminf_{a\to\infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in N) \geqq -H_\mu(Q). \tag{4.1}$$

Proof. For any $a \in \mathbf{Z}^d_+$ and for any $j \in \Lambda(a)$ define

$$\Delta_j^a = \{z \in \mathbf{Z}^d_<;\ z+j \in \Lambda(a),\ z+j<j\}. \tag{4.2}$$

Let $q(d\omega(0)|F_{\Delta_j^a})$ be the Q-conditional probability distribution of $\omega(0)$ given $F_{\Delta_j^a}$. From the assumption that $H_\mu(Q)$ is finite we have that $q(d\omega(0)|F_{\Delta_j^a})$ is absolutely continuous with respect to $\mu(d\omega(0))$ (Q-a.s.). Then define ψ_j^a the corresponding density.

Let $D(a)=\{\omega\in\Omega;\, R_{\Lambda(a),\omega}\in N\}$. Denote $\tilde{\Lambda}(a)=\Lambda(a)\setminus\{0\}$, then for $x\in X$ using the θ-invariance of Q we have

$$E_0^{\mu,x}(D(a))\geqq \int_{D(a)} \exp\Big[-\sum_{j\in\tilde{\Lambda}(a)} \log(\psi_j^a(\theta_j\omega))\Big]\,\delta_x(d\omega(0))\,q_0(d\omega_{\tilde{\Lambda}(a)}) \tag{4.3}$$

where $q_0(d\omega_{\tilde{\Lambda}(a)})$ and $E_0^{\mu,x}$ are respectively the Q-conditional expectation and the P_μ-conditional expectation of $\omega_{\tilde{\Lambda}(a)}$ given $F_{\{0\}}$.

In Appendix A.1 it is shown that the condition $H_\mu(Q)<\infty$ implies the convergence in $L^1(Q)$ of

$$|\Lambda(a)|^{-1}\sum_{j\in\tilde{\Lambda}(a)} \log[\psi_j^a(\theta_j\omega)]\to H_\mu(Q). \tag{4.4}$$

For $\varepsilon>0$ and $a\in\mathbf{Z}_+^d$ define

$$F(a)=\Big\{\omega;\,\Big||\Lambda(a)|^{-1}\sum_{j\in\tilde{\Lambda}(a)} \log[\psi_j^a(\theta_j\omega)]-H_\mu(Q)\Big|\leqq\varepsilon\Big\}.$$

Then from (4.3)

$$E_0^{\mu,x}(D(a))\geqq\exp-[|\Lambda(a)|(H_\mu(Q)+\varepsilon)]\int_{D(a)\cap F(a)} \delta_x(d\omega(0))\,q_0(d\omega_{\tilde{\Lambda}(a)}). \tag{4.5}$$

For $a\in\mathbf{Z}_+^d$ and $x\in X$ define

$$\phi(a,x)=\{E_0^{\mu,x}(D(a))\exp[|\Lambda(a)|(H_\mu(Q)+\varepsilon)]\}\wedge 1. \tag{4.6}$$

Then by (4.5) integrating respect to Q

$$\int\phi(a,\omega(0))\,Q_0(d\omega(0))\geqq Q(D(a)\cap F(a)). \tag{4.7}$$

By the ergodic theorem (and (4.4) which is a consequence too)

$$\lim_{a\to\infty} Q(D(a)\cap F(a))=1 \tag{4.8}$$

and then

$$\lim_{a\to\infty}\int\phi(a,\omega(0))\,Q_0(d\omega(0))=1. \tag{4.9}$$

But the condition $H_\mu(Q)<\infty$ implies also that Q_0 is absolutely continuous to μ and then there exists a constant $c>0$ such that

$$\liminf_{a\to\infty}\int\phi(a,x)\,\mu(dx)\geqq c \tag{4.10}$$

and then for any $\varepsilon>0$

$$\liminf_{a\to\infty}|\Lambda(a)|^{-1}\log P_\mu(D(a))\geqq -H_\mu(Q)+\varepsilon \tag{4.11}$$

and for $\varepsilon\to 0$ we have (4.1). □

Theorem 4.2. *Let $Q \in M_\theta(\Omega)$ be such that $H_\mu(Q) < \infty$. Then for any open neighborhood N in $M_\theta(\Omega)$*

$$\liminf_{a \to \infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in N) \geqq -H_\mu(Q). \tag{4.12}$$

Thus if A is any open set in $M_\theta(\Omega)$

$$\liminf_{a \to \infty} |\Lambda(a)|^{-1} \log P_\mu(R_{\Lambda(a),\omega} \in A) \leqq -\inf_{Q \in A} H_\mu(Q). \tag{4.13}$$

Proof. By the property (2.4) of $H_\mu(Q)$ it is enough to prove it for Q a finite convex combination of ergodic measures. Let then

$$Q = \sum_{h=0}^{I} p_h \cdot Q_h, \qquad \sum_{h=0}^{I} p_h = 1, \qquad Q_h \text{ ergodic.}$$

Divide $\Lambda(a)$ in disjoint sets $\Lambda_h(a)$, $h = 1, \ldots, I$, such that $|\Lambda_h(a)| = [p_h |\Lambda(a)|]$ (where $[x]$ means the integer part of x). Then the theorem follows by an easy modification of the argument used in [5 (IV)] Theorem 5.5. □

5. The Gibbs Variational Principle and Large Deviations for Gibbs Measures

Let X be a compact polish space and μ a probability measure on X. Let Φ be a translation invariant interaction, i.e. a family of continuous functions

$$\{\Phi_\Lambda : (\Omega, F_\Lambda) \to \mathbf{R}, \Lambda \subset \mathbf{Z}^d \text{ bounded}\} \quad \text{such that } \Phi_\Lambda(\theta_j \omega) = \Phi_{\Lambda+j}(\omega)$$
$$\text{for any bounded } \Lambda \subset \mathbf{Z}^d \text{ and any } j \in \mathbf{Z}^d.$$

An interaction Φ is of finite range if there exists a finite set $\Delta_\Phi \subset \mathbf{Z}^d$ such that if $0 \in \Lambda$ then $\Phi_\Lambda = 0$ if $\Delta_\Phi \not\supset \Lambda$.

Denote by A the Banach space of the translation invariant interactions endowed with the norm

$$|\Phi| = \sum_{\Lambda \ni 0} \frac{1}{|\Lambda|} \|\Phi_\Lambda\|_\infty \tag{5.1}$$

and let A_0 be the dense linear subspace of the finite range interactions (cf. [12]).

Define, for a $\Phi \in A$ and $\lambda \subset \mathbf{Z}^d$ bounded,

$$U_\Lambda^\Phi(\omega) = \sum_{X \subset \Lambda} \Phi_X(\omega) \tag{5.2}$$

$$Z_\Lambda^\Phi = \int_\Omega \exp[-U_\Lambda^\Phi(\omega)]\, P_\mu(d\omega) \tag{5.3}$$

where P_μ is the homogeneous product measure with marginal μ considered in the preceding sections.

For any $\Phi \in A$ define a continuous function $\mathrm{A}^{\Phi} \in C(\Omega)$ as

$$\mathrm{A}^{\Phi}(\omega) = -\sum_{\Lambda \ni 0} \frac{1}{|\Lambda|} \Phi_{\Lambda}(\omega). \tag{5.4}$$

The map $\Phi \in A \to \mathrm{A}^{\Phi} \in C(\Omega)$ is linear and continuous (because $\|\mathrm{A}^{\Phi}\|_{\infty} \leqq |\Phi|$).

Theorem 5.1 (The Gibbs Variational Formula). *Let $\Phi \in A$, then*

$$\lim_{a \to \infty} \frac{1}{|\Lambda(a)|} \log Z^{\Phi}_{\Lambda(a)} = \sup_{Q \in M_{\theta}(\Omega)} \{Q(\mathrm{A}^{\Phi}) - H_{\mu}(Q)\}. \tag{5.5}$$

Proof. From the large deviation result of Sects. 3 and 4 and Theorem 2.2 of [15]

$$\lim_{a \to \infty} \frac{1}{|\Lambda(a)|} \log \int_{\Omega} \exp\Big(\sum_{j \in \Lambda(a)} [\mathrm{A}^{\Phi}(\theta_j \omega^a)]\Big) dP_{\mu} = \sup_{Q \in M_{\theta}(\Omega)} \{Q(\mathrm{A}^{\Phi}) - H_{\mu}(Q)\}. \tag{5.6}$$

Then it is enough to show that the limit at the left hand side of (5.6) is equal to the left hand side of (5.5). We proceed as in ([12]; Theorem 3.4).

Suppose $\Phi \in A_0$, then for any $\omega \in \Omega$

$$\Big| U^{\Phi}_{\Lambda(a)}(\omega) + \sum_{j \in \Lambda(a)} \mathrm{A}^{\Phi}(\theta_j \omega^a) \Big| \leqq N(\Lambda(a)) |\Phi| \tag{5.7}$$

where $N(\Lambda(a)) = \{j \in \Lambda(a);\ \Lambda_{\Phi} + j \not\subset \Lambda(a)\}$, and then

$$\frac{N(\Lambda(a))}{|\Lambda(a)|} \to 0 \quad \text{as } a \to \infty.$$

It follows that, for $\Phi \in A_0$,

$$\lim_{a \to \infty} \frac{1}{|\Lambda(a)|} \Big\{ \log Z^{\Phi}_{\Lambda(a)} - \log \int_{\Omega} \exp\Big[\sum_{j \in \Lambda(a)} \mathrm{A}^{\Phi}(\theta_j \omega^a)\Big] dP_{\mu} \Big\} = 0. \tag{5.8}$$

Denote $P^{\Phi}_{\Lambda} = \frac{1}{|\Lambda|} \log Z^{\Phi}_{\Lambda}$; and for $\mathrm{A} \in C(\Omega)$ define

$$P_{\Lambda(a)}(\mathrm{A}) = \frac{1}{|\Lambda(a)|} \log \int_{\Omega} \exp\Big[\sum_{j \in \Lambda(a)} \mathrm{A}(\theta_j \omega^a)\Big] P_{\mu}(d\omega). \tag{5.9}$$

If A, $\mathrm{B} \in C(\Omega)$ it is easy to show that for any $a \in \mathbf{Z}^d_+$

$$|P_{\Lambda(a)}(\mathrm{A}) - P_{\Lambda(a)}(\mathrm{B})| \leqq \|\mathrm{A} - \mathrm{B}\|_{\infty}. \tag{5.10}$$

Then by the density of $\{\mathrm{A}^{\Phi};\ \Phi \in A_0\}$ in $C(\Omega)$ and the equicontinuity property (5.10), (5.5) follows for any $\Phi \in A$. □

The theorem above is true also if we consider fixed boundary conditions, i.e. for $\omega' \in \Omega$ let

$$Z_\Lambda^{\Phi *}(\omega') = \int \exp - [U_\Lambda^\Phi(\omega) + W_\Lambda^\Phi(\omega, \omega')] \, P_\mu(d\omega) \tag{5.11}$$

where

$$W_\Lambda^\Phi(\omega, \omega') = \sum_{X \cap \Lambda \neq \phi;\, X \cap \Lambda^c \neq \phi} \Phi_X(\omega_\Lambda \vee \omega'_{\Lambda^c}) \tag{5.12}$$

and if the interaction Φ satisfies the condition

$$\|\Phi\| = \sum_{\Lambda \ni 0} \|\Phi_\Lambda\|_\infty < +\infty. \tag{5.13}$$

In (5.12) we use the notation

$$\omega_\Lambda \vee \omega'_{\Lambda^c}(j) = \omega(j) \qquad \text{for } j \in \Lambda$$

$$\omega_\Lambda \vee \omega'_{\Lambda^c}(j) = \omega'(j) \qquad \text{for } j \in \Lambda^c.$$

Then we have, uniformly for any $\omega' \in \Omega$,

$$\lim_{a \to \infty} \frac{1}{|\Lambda(a)|} \log Z_{\Lambda(a)}^{\Phi *}(\omega') = \sup_{Q \in M_\theta(\Omega)} \{Q(\mathrm{A}^\Phi) - H_\mu(Q)\}. \tag{5.14}$$

The proof of (5.14) use the same argument of the proof of Theorem 5.1 (cf. [12]). In the following we will denote $P(\Phi)$ the R.H.S. of (5.14).

For $\Lambda \subset \mathbf{Z}^d$ bounded and $\omega \in \Omega$, define the Gibbs conditional probability distribution on Λ given ω' on Λ^c as

$$\pi_{\Lambda,\omega'}^\Phi(d\omega_\Lambda) = [Z_\Lambda^{\Phi *}(\omega')]^{-1} \exp - [U_\Lambda^\Phi(\omega) + W_\Lambda^\Phi(\omega, \omega')] \, \mu_\Lambda(d\omega_\Lambda). \tag{5.15}$$

A probability measure ν on Ω is a Gibbs measure for the interaction $\Phi \in A$ if for any $\Lambda \subset \mathbf{Z}^d$ bounded and any $f \in C(\Omega)$

$$\nu(f) = \int \pi_{\Lambda,\omega'}^\Phi(f) \, \nu(d\omega'). \tag{5.16}$$

The Eq. (5.16) are called DLR equations (cf. [9, 12]). Denote with G^Φ the set of all Gibbs measure for the interaction Φ. The following theorem establish a uniform large deviations principle for the Gibbs conditional distributions defined by (5.15).

Theorem 5.2. *For any $\omega' \in \Omega$, any $C \subset M_\theta(\Omega)$ closed and any $G \subset M_\theta(\Omega)$ open*

$$\limsup_{a \to \infty} \frac{1}{|\Lambda(a)|} \log \pi_{\Lambda,\omega'}^\Phi(R_{\Lambda(a)} \in C) \leqq -\inf_{Q \in C} H_\mu^\Phi(Q) \tag{5.17}$$

$$\liminf_{a \to \infty} \frac{1}{|\Lambda(a)|} \log \pi_{\Lambda,\omega'}^\Phi(R_{\Lambda(a)} \in G) \geqq -\inf_{Q \in G} H_\mu^\Phi(Q) \tag{5.18}$$

where

$$H_\mu^\Phi(Q) = H_\mu(Q) - Q(\mathrm{A}^\Phi) + P(\Phi). \tag{5.19}$$

Proof. Again using Theorem 2.2 of [15] it is possible to show that for any $C \subset M_\theta(\Omega)$ closed

$$\limsup_{a\to\infty} \frac{1}{|\Lambda(a)|} \log \int_{\{R_{\Lambda(a)}\in C\}} \exp\left[\sum_{j\in\Lambda(a)} A^{\Phi}(\theta_j\omega^a)\right] P_\mu(d\omega)$$
$$\leqq \sup_{Q\in C}\{Q(A^{\Phi})-H_\mu(Q)\}. \tag{5.20}$$

Then by (5.14)

$$\limsup_{a\to\infty} \frac{1}{|\Lambda(a)|} \log \pi^{\Phi}_{\Lambda(a),\omega'}(R_{\Lambda(a)}\in C)$$
$$= \limsup_{a\to\infty} \frac{1}{|\Lambda(a)|} \log \int_{\{R_{\Lambda(a)}\in C\}} \exp-[U^{\Phi}_{\Lambda}(\omega)+W^{\Phi}_{\Lambda}(\omega,\omega')]\, P_\mu(d\omega) - P(\Phi)$$
$$\leqq \sup_{Q\in C}\{Q(A^{\Phi})-H_\mu(Q)\} - P(\Phi) \tag{5.21}$$

where in the last inequality we have used the same density argument of the proof of Theorem 5.1. The proof of (5.18) is analogous. □

The next theorem gives the large deviation principle for any Gibbs measure for a given interaction $\Phi\in A$ and its proof is an immediate consequence of Theorem 5.2, the DLR equations (5.16) and the continuity of the map $\omega\in\Omega\to\pi^{\Phi}_{\Lambda,\omega}$.

Observe that the result depends only on the interaction Φ and the reference measure μ and not on the particular Gibbs measure ν chosen in G^{Φ}. This fact follows from the uniformity of the estimate for the Gibbs conditional distributions in the boundary conditions (cf. Theorem 5.2).

Theorem 5.3. *Given an interaction $\Phi\in A$, then for any $\nu\in G^{\Phi}$, for any closed set C and any open set $G\in M_\theta(\Omega)$*

$$\limsup_{a\to\infty} \frac{1}{|\Lambda(a)|} \log \nu(R_{\Lambda(a)}\in C) \leqq -\inf_{Q\in C} H^{\Phi}_\mu(Q) \tag{5.22}$$

$$\liminf_{a\to\infty} \frac{1}{|\Lambda(a)|} \log \nu(R_{\Lambda(a)}\in G) \leqq -\inf_{Q\in G} H^{\Phi}_\mu(Q) \tag{5.23}$$

where $H^{\Phi}_\mu(Q)$ is defined by (5.18).

6. The Lanford Theory

In this section we deduce the Lanford theory of large deviations for finite range observables (cf. [9]), using the results obtained for the empirical distribution. In fact the Theorem 5.3 gives us all the information that we need to control the large deviations for the observables. The tool used to obtain this information is the so called "contraction principle" (cf. [15]).

Let us first give the definition of finite range observable. As in Sect. 5 we assume that X is compact and that the reference measure μ on X is normalized.

Definition. A finite range observable is a set of continuous functions $\{f_\Lambda: \Omega\to\mathbf{R}, \Lambda\subset\mathbf{Z}^d \text{ finite}\}$ such that:

a) f_Λ is F_Λ-measurable

b) $f_{\Lambda+j}=f_\Lambda\circ\theta_j \; j\in\mathbf{Z}^d$

c) $f_{\Lambda\cup\Lambda'}=f_\Lambda+f_{\Lambda'}$ if the distance between Λ and Λ' is larger than $r(f)$ (that we call the range of the observable).

d) $f_\phi=0$.

Given an observable $\{f_\Lambda\}$, one can define recursively a translation invariant finite range interaction $\Psi^f\in A_0$ (as in the definition of Sect. 5) such that

$$f_\Lambda=\sum_{X\subset\Lambda}\Psi_X^f \tag{6.1}$$

It is also easy to check that given a finite range interaction $\Psi\in A_0$, (6.1) define a finite range observable in the sense of the definition above, with $r(f)$ given by the maximum diameter of the range of Ψ^f. So we have a $1-1$ correspondence between observables and translation invariant interactions. Then as in Sect. 5, for a given observable define the continuous function on Ω

$$\mathrm{A}^{\mathrm{f}}(\omega)=-\sum_{\Lambda\ni 0}\frac{1}{|\Lambda|}\Psi_\Lambda^f(\omega). \tag{6.2}$$

Theorem 6.1. *Let $\Phi\in A$ satisfying* (5.13) *and f_Λ be a finite range observable, then for any $v^\Phi\in G^\Phi$:*

a) *if $J\subset\mathbf{R}$ is closed*

$$\limsup_{a\to\infty}\frac{1}{|\Lambda(a)|}\log v^\Phi(|\Lambda(a)|^{-1}f_{\Lambda(a)}\in J)\leqq -\inf_{x\in J}I(x;f,\Phi) \tag{6.3}$$

b) *if $J\subset\mathbf{R}$ is open*

$$\liminf_{a\to\infty}\frac{1}{|\Lambda(a)|}\log v^\Phi(|\Lambda(a)|^{-1}f_{\Lambda(a)}\in J)\geqq -\inf_{x\in J}I(x;f,\Phi) \tag{6.4}$$

where

$$I(x;f,\Phi)=\inf\{H_\mu^\Phi(Q);\; Q\in M_\theta(\Omega), Q(\mathrm{A}^f)=x\} \tag{6.5^1}$$

or

$$I(x;f,\Phi)=+\infty \quad \text{if} \quad \{Q\in M_\theta(\Omega), Q(\mathrm{A}^f)=x\}=\phi. \tag{6.5^2}$$

Remark. $I(x;f;\Phi)=+\infty$ if $|x|>|\mathrm{A}^f|_\infty$, i.e., the set $\{x\in\mathbf{R}, I(x)<\infty\}$ is relatively compact

Proof. Define

$$S_{\Lambda(a)}(\omega)=R_{\Lambda(a),\omega}(\mathrm{A}^f)=|\Lambda(a)|^{-1}\sum_{j\in\Lambda(a)}\mathrm{A}^f(\theta_j\omega^a) \tag{6.6}$$

and let be γ_a the distribution of $S_{\Lambda(a)}(\omega)$ on $\mathbf{R}$ induced by v^Φ, i.e., for $J\subset\mathbf{R}$

$$\gamma_a(J)=v^\Phi(S_{\Lambda(a)}\in J). \tag{6.7}$$

Observe that the map $Q \to Q(A^f)$ is continuous from $M_\theta(\Omega)$ to $\mathbf{R}$. Then by Theorem 5.3 and the contraction principle (cf. [15]) the family of probability measures $\{\gamma_a, a \in \mathbf{Z}^d\}$ satisfies a large deviations principle with rate function $I(x; f, \Phi)$.

Because the range of f is finite, we have for any $\omega \in \Omega$

$$\left| S_{\Lambda(a)}(\omega) - \frac{1}{|\Lambda(a)|} f_{\Lambda(a)}(\omega) \right| \leqq \frac{N_f(\Lambda(a))}{|\Lambda(a)|} \tag{6.8}$$

where

$$N_f(\Lambda(a)) = \{ j \in \Lambda(a);\ \Delta_{\Psi^f} + j \not\subset \Lambda(a) \}.$$

By the affinity and the lower semicontinuity of H_μ^Φ we have that $I(x; f, \Phi)$ is convex and lower semicontinuous. For a closed set J define

$$J^2 = \left\{ x \in \mathbf{R};\ \mathrm{dist}(x, J) \leqq \frac{N_f(\Lambda(a))}{|\Lambda(a)|} \|\Psi^f\| \right\}, \tag{6.9}$$

then $J^2 \downarrow J$ if $a \to \infty$. For $\varepsilon > 0$ let a' such that

$$\inf_{x \in J^{a'}} I(x; f, \Phi) \geqq \inf_{x \in J} I(x; f, \Phi) - \varepsilon. \tag{6.10}$$

Let $a \in \mathbf{Z}_+^d$ such that $\Lambda(a) \supset \Lambda(a')$, then

$$\nu^\Phi(|\Lambda(a)|^{-1} f_{\Lambda(a)} \in J) \leqq \gamma_a(J^a) \leqq \gamma_a(J^{a'})).$$

Then for any $\varepsilon > 0$ we have

$$\limsup_{a \to \infty} \frac{1}{|\Lambda(a)|} \log \nu^\Phi(|\Lambda(a)|^{-1} f_{\Lambda(a)} \in J) \leqq -\inf_{x \in J} I(x; f, \Phi) + \varepsilon$$

and if we let $\varepsilon \to 0$ we obtain (6.3).

The proof of (6.4) is analogous considering, for a given open $J \subset \mathbf{R}$, the sets

$$J^a = \left\{ x \in J;\ \mathrm{dist}(x, J^c) > \frac{N_f(\Lambda(a))}{|\Lambda(a)|} \|\Psi^f\| \right\}. \quad \square$$

Theorem 6.1 extends immediately to $\mathbf{R}^n$-valued observables, for any $n \in \mathbf{N}$.

Remark. By the lower semicontinuity of H_μ^Φ and the continuity of A^f, it follows that $I(x; f, \Phi) = 0$ if only if there exists a translation invariant Gibbs measure $\nu \in G^\Phi \cap M_\theta(\Omega)$ such that $\nu(A^f) = x$. Thus $I(x; f; \Phi)$ may have more than one zero only if there is a phase transition. Conversely it is sufficient to find an observable with a corresponding I-function with a non unique zero in order to establish the existence of more than one translation invariant Gibbs measure.

Appendix

We prove here, using a result of Barron (cf. [3]), the formula (4.4) used in the proof of the lower estimate (Sect. 4).

Lemma A 1. *Under the assumptions of Lemma* 3.1

$$\lim_{a\to\infty} |\Lambda(a)|^{-1} \sum_{j\in\tilde{\Lambda}(a)} \log(\psi_j^a \theta_j) = H_\mu(Q) \tag{A.1}$$

Q-almost surely and in $L^1(Q)$.

Proof. Let Γ_n any increasing sequence of subsets of $\mathbf{Z}^d_<$ such that $\Gamma_n \uparrow \mathbf{Z}^d_<$. Then by [3], Lemma 2, $\{\log \psi_{\Gamma_n}\}$ is bounded in $L^1(Q)$ and

$$\log \psi_{\Gamma_n} \to \log \psi_< \quad \text{in } L^1(Q) \text{ and } Q\text{-a.s.} \tag{A.1}$$

Let Δ_j^a the sequence of sets defined by (4.2). Then consider $\Gamma_n = \{z \in \mathbf{Z}^d_<; |z^i| \leqq n; i=1, \ldots, d\}$. For any $\varepsilon > 0$ there exists n enough large such that if $\Gamma_n \subset \Delta_j^a$, then

$$\|\log \psi_{\Delta_j^a} - \log \psi_<\|_{L^1(q) \leqq \varepsilon}.$$

Then

$$\sum_{j\in\Lambda(a)} \|\log \psi_{\Delta_j^a} \cdot \theta_j - \log \psi_< \cdot \theta_j\|_{L^1(q)} \leqq \varepsilon |\Lambda(a)| + |\Lambda(a)^*| \sup_{\Gamma_n} \|\log \psi_{\Gamma_n}\|_{L^1(Q)}$$

where $\Lambda(a)^* = \{j \in \Lambda(a); \Gamma_n \not\subset \Delta_j^a\}$.

We have that

$$\frac{|\Lambda(a)^*|}{|\Lambda(a)|} \to 0 \quad \text{as } a \to \infty$$

and (A.1) follows from the arbitrarity of ε and the d-dimensional ergodic theorem. □

Acknowledgements. The author thanks S.R.S. Varadhan for the helpful discussions about large deviations. A special thanks to J. Lebowitz, R. Shonmann, and J. Fritz for their interest, encouragement, and suggestions. Thanks also to S. Orey and to the Institute of Mathematics of the University of Minnesota for their kind hospitality. The work was partially supported by NSF-DMR81-14726.

References

1. Accardi, L., Olla, S.: Donsker and Varadhan's theory for stationary processes. Preprint June 1984
2. Baldi, P.: Large deviations and stochastic homogenization. Preprint 1985
3. Barron, A.R.: The strong ergodic theorem for densities: generalized Shannon McMillan-Breiman theory. Ann. Probab. **13**, 4 (1985)
4. Dawson, D.A., Gartner, J.: Large deviations from the McKean-Vaslov limit for weakly interacting diffusions, Stochastics 20, 247–308 (1987)
5. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time (I), Comm. Pure Appl. Math. **28**, 1–47 (1975); (IV) Commun. Pure Appl. Math. **36**, 182–212 (1983)

6. Donsker, M.D., Varadhan, S.R.S.: Large deviations for stationary gaussian processes. Commun. Math. Phys. **97**, 187–210 (1985)
7. Föllmer, H.: Entropy and information gain in random fields. Z. Wahrscheinlichkeitstheor. Verw. Geb. **20**, 207–217 (1973)
8. Ellis, R.S.: Entropy, large deviations and statistical mechanics. (Grundlehren der Mathematischen Wissenschaften vol. 271) Berlin Heidelberg New York: Springer 1985
9. Lanford, O.E.: Entropy and equilibrium states in classical statistical mechanics. (Lect. Notes Phys. vol. 20) Berlin Heidelberg New York: Springer 1973
10. Lebowitz, J.L., Schonmann, R.H.: Pseudo free energies and large deviations for nongibbsian FKG measures. Preprint 1986
11. Olla, S.: Large deviations for almost markovian processes. Prob. Th. Rel. Fields **76**, 395–409 (1987)
12. Ruelle, D.: Thermodynamical formalism, Encyclopedia of mathematics and its applications, vol. 5. Reading Mass.: Addison Wesley 1978
13. Strook, D.W.: An Introduction to the Theory of Large Deviations. Berlin Heidelberg New York: Springer 1984
14. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261–296 (1966)
15. Varadhan, S.R.S.: Large deviations and Applications, (CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 46) Philadelphia: Society for Industrial and Applied Mathematics 1984
16. Comets, F.: Grandes deviations pour des champs de Gibbs sur Z^d. Preprint 1986
17. Föllmer, H., Orey, S.: Large deviations for the empirical field of a Gibbs measure. IMA Preprint Series No. 257, August 1986

Received June 6, 1986; received in revised form October 20, 1987