276

Probab. Th. Rel. Fields 77, 49–64 (1988)

Probability Theory and Related Fields

Pseudo-Free Energies and Large Deviations for Non Gibbsian FKG Measures*

J.L. Lebowitz** and R.H. Schonmann***
Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Summary. A large deviation theorem for the invariant measures of translation invariant attractive interacting particle systems on $\{0, 1\}^{Z^d}$ is proven. In this way a pseudo-free energy and pressure is defined. For ergodic systems the large deviations property holds with the usual scaling. The case of non ergodic systems is also discussed. A similar result holds for occupation times. The perturbation by an external field is treated.

1. Introduction

Gibbs measures are known to describe the properties of macroscopic physical systems in equilibrium. For a system in a box $\Lambda \subset R^d$ with microscopic interaction U_Λ, in equilibrium at temperature β^{-1}, the appropriate statistical state is according to Gibbs and Einstein $\mu_{\Lambda,\beta} \sim \exp(-\beta U_\Lambda)$ [Ru1], [Ru2], [Sin]. The structure of the infinite volume limit $\Lambda \to R^d$, necessary for making sharp statements about macroscopic phenomena, such as phase transitions, has been much studied and a lot is known about them. This is particularly so when the physical object to be represented can be modeled (and this happens surprisingly often) as a "spin" system on a lattice with "sufficiently rapidly" decaying interactions U. U_Λ is just U restricted to Λ plus boundary terms. In this case the $\mu_\beta = \lim_{\Lambda \to Z^d} \mu_{\Lambda,\beta}$ are quasi-Markovian measures on the compact configuration space $E_d = W^{Z^d}$, $d = 1, 2, \ldots$, $W = (w_1, \ldots, w_r)$, $w_1 < \ldots < w_r$, i.e. their conditional probabilities in a finite region $\Lambda \subset Z^d$, specified by the DLR (Dobrushin-Lanford-Ruelle) equations [Ru1], [Ru2], [Sin], depend only weakly (or not at all) on what the configuration is far away from Λ. They are in fact just the finite volume states $\mu_{\Lambda,\beta}$ with suitable boundary conditions. This "locality" captures the essence

* Work partially supported by NSF-DMR81-14726 (USA) and CNPq (Brazil)
** Also Department of Physics
*** *Permanent address:* Mathematics Department, University of São Paulo, Caixa Postal 20570, 01000 São Paulo, SP Brazil

of the equilibrium state of macroscopic physical systems – if we divide the system into two or more parts and isolate them from each other then the separate regions of the system continue to be in the same equilibrium state as before.

The situation is very different when we consider the behavior of systems maintained in a nonequilibrium state by contacts with outside sources. For such systems the appropriate representations, i.e. measures which can be used to obtain the properties of stationary non-equilibrium states, cannot be expected to behave in a quasi-Markovian way – isolating a part will generally change its behavior drastically. This makes the study of such measures more difficult and only very little is known at the present time about them or even how to characterize them in a generally useful way [G.K.I.], [G.L.P.], [K.L.S.], [Leb2]. In fact there may not be any general formalism comparable to the Gibbsian equilibrium one which will encompass the great variety of nonequilibrium behavior observed in nature, even when restricted to steady state situations. Nevertheless the subject is clearly of great interest and in this paper we study some aspects of simple models of such measures.

Before going on to describe our new work let us review very briefly some of the features of Gibbs measures which may, or may not, be generalizable to non-equilibrium systems. We refer the reader to articles by Gray [Gra], Künsch [Kün] and references there for discussions and some results on this question. It follows from the quasi-Markovian nature of the DLR equations that all Gibbs measures are obtained as infinite volume limits of finite volume measures $\mu_{\Lambda,\beta} \sim \exp(-\beta U_\Lambda)$ with suitable boundary conditions. (These b.c. may have to be statistical – but in all known cases can be taken pure, i.e. there is a specified configuration on sites outside Λ). The states form a Choquet simplex whose extremal points generally have rapidly decaying correlations, exponential if the interactions are finite range, except at places where there are good reasons why they shouldn't, i.e. at "critical" points or lines.

The translational invariant states are characterized by a variational principle – their extremal points are the pure phases and correspond to tangent planes of a convex functional, the pressure p on an appropriate Banach space of potentials βU [Ru1], [Ru2], [Sin]. First order phase transitions occur at values of βU for which the tangent plane is not unique – so that there is more than one extremal translation invariant state corresponding to a coexistence of pure phases. For a given physical systems the potential U determining the Gibbs measure usually contains one or more parameters which can be varied experimentally or at least can be imagined so. The phase diagram of the system is a picture of how the number of pure phases changes when these parameters are changed.

The most important of these parameters is the magnetic field h in spin language (chemical potential in particle language) which controls the magnetization (or density). Considered as a function of h, $p(h)$ is convex and its derivative (which exists for almost all h) gives the average magnetization. Its second derivative is intimately related to the variance of the fluctuations in the magnetization. It also contains information about large deviations from the average (those proportional to the volume). In this way it describes (for some ferromagnetic systems it gives complete information) the coexistence of pure phases [Ru1], [Ru2], [Sin].

The present paper may be thought of as an attempt, with some measure of success, to generalize this aspect of Gibbs states to some nonequilibrium states. The states we have particularly in mind are stationary measures for stochastic time evolutions of infinite particle systems on a lattice. These type of evolutions have been investigated in recent years from various points of view [Lig]. Only little is known however about the global structure of the stationary states even in the simplest examples which are not explicitely constructed to be Gibbs. There are however many cases where these states are known to satisfy the FKG (Fortuin, Kasteleyn, Ginibre) inequalities [F.K.G.], i.e. for any increasing and continuous $f, g: E_d \to R$. $E_\nu(f(\eta)g(\eta)) \geqq E_\nu(f(\eta))E_\nu(g(\eta))$ where E_ν is the expectation when η is random with law ν. In the terminology of [Lig] this measures are said to have *positive correlations*. In this note we show that these states share at least some features of Gibbs states. In particular it is possible to define a pressure like function $\Pi(h)$ which is related to the large fluctuations in the invariant state ν in a manner similar to that of $p(h)$ in equilibrium systems. It reduces to it (up to location of the origin of h) for Gibbs measures.

Statement of Results

Before stating the precise results we must introduce some notation. For $\eta \in E_d$ let $\eta(i)$ represent the state at site $i \in Z^d$. Given a set S, $|S|$ will be its cardinality. For $\Lambda \subset Z^d$ and $\eta \in E_d$ define

$$X_\Lambda(\eta) = |\Lambda|^{-1} \sum_{i\in\Lambda} \eta(i).$$

For simplicity we will write $\nu\{\eta(0)=1\}$ instead of $\nu\{\eta: \eta(0)=1\}$ and $\nu\{X_\Lambda \geqq x\}$ instead of $\nu\{\eta: X_\Lambda(\eta) \geqq x\}$, etc.

In Sect. 2 we will prove

Theorem 1. *Consider a probability measure ν on E_d which is translation invariant,* FKG *and such that $\rho_i = \nu\{\eta(0)=w_i\} > 0$, $i=1, r$. Let $(\Lambda_n)=(\Lambda)$ be a sequence of cubes in Z^d such that $\Lambda_n \to Z^d$. Then,*

a) for any $x \in [w_1, w_r]$, $|\Lambda|^{-1} \log \nu\{X_\Lambda \geqq x\}$ [resp. $|\Lambda|^{-1} \log \nu\{X_\Lambda \leqq x\}$] *converges as $\Lambda \to Z^d$ to a non positive real valued function $\lambda_+(x)$* [resp. $\lambda_-(x)$] *which is concave and decreasing* [resp. increasing], *$\lambda_+(w_1)=0$ and $\lambda_+(w_r) \geqq \log \rho_r > -\infty$* [resp. $\lambda_-(w_1) \geqq \log \rho_1 > -\infty$ *and* $\lambda_-(w_r)=0$].

b) Define $\lambda(x)=\min(\lambda_-(x), \lambda_+(x))$, then $\lambda: [w_1, w_r] \to (-\infty, 0]$ is concave and for any $w_1 \leqq a < b \leqq w_r$ such that

$$\min(\lambda(a), \lambda(b)) < 0 \tag{1.1}$$

$$\lim_{\Lambda\to Z^d} |\Lambda|^{-1} \log \nu\{X_\Lambda \in J\} = \sup_{a \leqq x \leqq b} \lambda(x) \tag{1.2}$$

for $J=[a, b]$, $[a, b)$, $(a, b]$, (a, b).

(1.3) *c)*
$$\lim_{\Lambda\to Z^d} |\Lambda|^{-1} \log E_\nu \exp(h|\Lambda| X_\Lambda) = \Pi(h)$$

where $\Pi: R \to R$ *is defined by*

(1.4) $$\Pi(h) = \sup_{w_1 \leqq x \leqq w_r} (\lambda(x) + h\,x).$$

In particular Π *is convex.*

Remark. If the condition $\rho_i > 0$, $i = 1, r$ is not true, one can modify the definition of W in order to make it hold.

Note that the restriction (1.1) is empty unless the set $\mathscr{F} = \{x \in [w_1, w_r]: \lambda(x) = 0\}$ has a positive width.

In Sect. 3 we investigate the invariant measures of some interacting particle systems. We consider in particular the class of translation invariant attractive spin systems [Lig], hereafter denoted by TIA. These are Markov and Feller processes with state space $\{0, 1\}^{Z^d}$ whose evolution are given by the flip rates $c(i, \eta)$ (the rate at which $\eta(i)$ flips to $1 - \eta(i)$ when the system is in the configuration η). Translation invariant means that $c(i, \eta) = c(i + j, \tau_j \eta)$, where $(\tau_j \eta)(k) = \eta(k + j)$. Attractiveness means informally that zeros attract zeros and ones attract ones ("ferromagnetic" types). More precisely, if the configuration η is dominated by the configuration ζ, i.e. $\eta(j) = 1 \Rightarrow \zeta(j) = 1, j \in Z^d$, then

$$c(i, \eta) \leqq c(i, \zeta) \quad \text{if } \eta(i) = \zeta(i) = 0$$

$$c(i, \eta) \geqq c(i, \zeta) \quad \text{if } \eta(i) = \zeta(i) = 1.$$

In order that the infinitesimal rates $c(i, \eta)$ define *a* unique process one must assume that they do not depend very strongly on the configurations at sites far away from i; a sufficient condition can be found in Chap. III of [Lig].

As in [Lig], we will denote by $S(t)$ the corresponding semigroup and write $\mu S(t)$ for its action on *a* measure.

Some of the fundamental facts about the TIA are summarized next (for proofs see [Lig]).

(1.5) $\delta_0 S(t)$ [resp. $\delta_1 S(t)$] converges weakly to a measure ν_- [resp. ν_+] which is invariant for $S(t)$ (δ_k is the point mass on the configuration $\eta(i) = k$ for all i).

(1.6) The process is ergodic iff $\nu_- = \nu_+$.

(1.7) ν_- and ν_+ are translation invariant and ergodic with respect to translations. They are also FKG.

(1.8) If μ is translation invariant and FKG and $\mu S(t) \to \nu$ weakly, then ν is invariant for $S(t)$ and is also translation invariant and FKG. Theorem 1 therefore applies to these measures and the next theorem gives information about the corresponding $\lambda(x)$.

Define $\rho_\pm = \nu_\pm \{\eta(0) = 1\}$.

Theorem 2. *Suppose that* ν *is an invariant measure for a* TIA. *Then there are constants* $C, \gamma > 0$, *which depend on* x, *such that*

(1.9) $$\textit{if } x < \rho_-, \quad \nu\{X_\Lambda \leqq x\} \leqq C\, e^{-\gamma|\Lambda|}$$

(1.10) $$\textit{if } x > \rho_+, \quad \nu\{X_\Lambda \geqq x\} \geqq C\, e^{-\gamma|\Lambda|}$$

It follows that

Corollary 1. *Suppose that ν above is also translation invariant and* FKG *and is neither δ_0 nor δ_1. Let $\lambda(x)$ be defined as in Theorem* 1, *then $\lambda(x)<0$ for $x<\rho_-$ or $x>\rho_+$.*

Remark. The hypothesis that $\nu \notin \{\delta_0, \delta_1\}$ is equivalent to the condition $\nu\{\eta(0)=0\} \neq 0$, $\nu\{\eta(0)=1\} \neq 0$, necessary to apply Theorem 1.

In particular if the system is ergodic the unique invariant measure is translation invariant and FKG by (1.5), (1.6) and (1.7). Therefore

Corollary 2. *Let ν be the unique invariant measure of an ergodic* TIA *and suppose $\nu \notin \{\delta_0\ \delta_1\}$. Then there exists a concave function λ: $[0,1] \to (-\infty, 0]$ such that $\{x \in [0,1]: \lambda(x)=0\} = \{\rho\} = \{\nu\{\eta(0)=1\}\}$ and* (1.2) *holds for any $0 \leqq a < b \leqq 1$.*

In Sect. 3 we also present some extensions of these results for more general increasing functions of the configuration and for occupation times. We consider some examples of TIA: the contact and voter models and finally discuss the relation between the large deviations and central limit theorems.

In Sect. 4 we consider the perturbation of a measure ν on E_d by an external field h in the following sense. Let ν_Λ be the measure induced by ν on W^Λ and define another measure $\nu_{\Lambda,h}$ on W^Λ by (here we use η to represent configurations on W^Λ)

(1.11) $$\nu_{\Lambda,h}(\eta) = (Z(\Lambda,h))^{-1}\, \nu_\Lambda(\eta) \exp(h\,|\Lambda|\, X_\Lambda)$$

where

(1.12) $$Z(\Lambda,h) = \sum_{\eta \in W^\Lambda} \nu_\Lambda(\eta) \exp(h\,|\Lambda|\, X_\Lambda) = E_\nu(\exp(h\,|\Lambda|\, X_\Lambda)).$$

Theorem 3. *If ν satisfies* (1.2) *for some concave $\lambda(x)$ and any $w_1 \leqq a < b \leqq w_r$, then*

(1.13) $$\lim_{\Lambda \to Z^d} |\Lambda|^{-1} \log \nu_{\Lambda,h}\{X_\Lambda \in [a,b]\} = \sup_{a \leqq x \leqq b} \lambda_h(x)$$

where

(1.14) $$\lambda_h(x) = \lambda(x) + x\,h - \Pi(h) = \inf_{h' \in R} (\Pi(h+h') - \Pi(h) - h'\,x).$$

It follows that if $\lambda(x)=0$ at a single point x, the family of measures $\nu_{\Lambda,h}$ does not show a "phase transition" (in the sense of a discontinuity of $x_{-,h} = \inf \mathscr{F}_h$ or $x_{+,h} = \sup \mathscr{F}_h$, where $\mathscr{F}_h = \{x \in [w_1, w_r]: \lambda(x)=0\}$) for small h.

In Sect. 5 we compare briefly our approach to the large deviations problem with other approaches.

2. Consequences of the FKG Relations

Proof of Theorem 1. It follows from FKG that if $\Gamma = \Gamma_1 \cup \Gamma_2$ then

(2.1) $$\nu\{(X_\Gamma \geqq x\} \geqq \nu\{X_{\Gamma_1} \geqq x\}\ \nu\{X_{\Gamma_2} \geqq x\}.$$

Define

$$A(n_1, \ldots, n_d) = \log \nu\{X_\Gamma \geqq x\}$$

where Γ is a rectangle of sides $n_1, \ldots, n_d$. Then from (2.1) and translation invariance

$$A(n_1, \ldots, n_k + n_k', \ldots, n_d) \geqq A(n_1, \ldots, n_k, \ldots, n_d) + A(n_1, \ldots, n_k', \ldots, n_d). \tag{2.2}$$

It now follows by a standard argument that $|\Lambda|^{-1} \log \nu\{X_\Lambda \geqq x\}$ converges. (See for instance the proof of step 1 of Theorem 2.6 of Chap. V of [Lig] for the case $d=1$, the general case is analogous.)

The facts that λ_+ is decreasing and $\lambda_+(0)=0$ are obvious. Using translation invariance and FKG again

$$\nu\{X_\Lambda \geqq w_r\} = \nu\{\eta(i) = w_r,\ i \in \Lambda\} \geqq \prod_{i \in \Lambda} \nu\{\eta(i) = w_r\} = (\rho_r)^{|\Lambda|}$$

which implies that $\lambda_+(w_r) \geqq \log \rho_r$.

Finally we will prove that λ_+ is concave, i.e. for any $x, y \in [w_1, w_r]$ and any $\alpha \in [0, 1]$,

$$\lambda_+(\alpha x + (1-\alpha) y) \geqq \alpha \lambda_+(x) + (1-\alpha) \lambda_+(y) \tag{2.3}$$

First we consider the case $\alpha = 1/2$. Take 2^d cubes of side n, $\Gamma_1, \Gamma_2, \ldots, \Gamma_{2^d}$ in Z^d such that their union is a cube Γ of side $2n$. By FKG

$$\begin{aligned} \nu\{X_\Gamma \geqq x/2 + y/2\} &\geqq \nu\{X_{\Gamma_i} \geqq x \text{ for } i=1, \ldots, 2^{d-1} \text{ and } X_{\Gamma_j} \geqq y \text{ for } j = 2^{d-1}+1, \ldots, 2^d\} \\ &\geqq \prod_{i=1}^{2^{d-1}} \nu\{X_{\Gamma_i} \geqq x\} \prod_{j=2^{d-1}+1}^{2^d} \nu\{X_{\Gamma_j} \geqq y\} \\ &= (\nu\{X_{\Gamma_1} \geqq x\})^{2^{d-1}} (\nu\{X_{\Gamma_1} \geqq y\})^{2^{d-1}} \end{aligned} \tag{2.4}$$

where we used translation invariance in the last equality. (2.3) with $\alpha = 1/2$ follows easily from (2.4). As is well known, by induction (2.1) follows then for any diadic rational α, i.e. $\alpha = p\,2^{-q}$ where p and q are integers. Finally we use the fact that λ_+ is decreasing to conclude. Suppose that $x<y$, take a sequence α_n of diadic rationals which increases and converges to α. Then for any n

$$\lambda_+(\alpha x + (1-\alpha) y) \geqq \lambda_+(\alpha_n x + (1-\alpha_n) y) \geqq \alpha_n \lambda_+(x) + (1-\alpha_n) \lambda_+(y).$$

Make $n \to \infty$ to conclude.

b) $\lambda(\cdot)$ is concave since it is the minimum of two concave functions. In particular it follows that it is continuous on (w_1, w_r). The proof of (1.2) is divided in many cases; we leave to the reader the cases $a = w_1$ or $b = w_r$ and consider only $w_1 < a < b < w_r$. In this case the statements are equivalent for the four types of intervals that J may represent; to see this fact just use relations like

$$\nu\{X_\Lambda\in[a,b-\varepsilon]\}\leqq\nu\{X_\Lambda\in[a,b)\}\leqq\nu\{X_\Lambda\in[a,b]\}\leqq\nu\{X_\Lambda\in[a,b+\varepsilon]\}$$

for $\varepsilon>0$. Suppose now that $\lambda(a)=0$, $\lambda(b)<0$. Then $\lambda(b)=\lambda_+(b)<0$ and $\lambda_+(a)=0$, so $\nu\{X_\Lambda\geqq b\}/\nu\{X_\Lambda\geqq a\}\to 0$ as $\Lambda\to Z^d$. Therefore

$$|\Lambda|^{-1}\log\nu\{X_\Lambda\in[a,b)\}=|\Lambda|^{-1}\log\nu\{X_\Lambda\geqq a\}$$
$$+|\Lambda|^{-1}\log(1-\nu\{X_\Lambda\geqq b\}/\nu\{X_\Lambda\geqq a\})\to\lambda_+(a)=0=\sup_{a\leqq x\leqq b}\lambda(x).$$

The case $\lambda(a)<0$, $\lambda(b)=0$ is analogous. Suppose $\lambda(a)<0$ and $\lambda(b)<0$; there are then in principle four possibilities: $\lambda(a)=\lambda_\alpha(a)<0$, $\lambda(b)=\lambda_\gamma(b)<0$, α, $\gamma=\pm$. The case $\alpha=+$, $\gamma=-$ is ruled out by the fact that for any $x\in[w_1,w_r]$,

$$\max(\lambda_-(x),\lambda_+(x))=0, \tag{2.5}$$

since otherwise $\nu\{X_\Lambda\leqq x\}+\nu\{X_\Lambda\geqq x\}\to 0$, which is absurd.

The cases $\alpha=\gamma$ are analogous, let us consider $\alpha=\gamma=+$: since λ_+ is concave and $\lambda_+(w_1)=0$, it follows that $\lambda_+(a)>\lambda_+(b)$ and therefore $\nu\{X_\Lambda\geqq b\}/\nu\{X_\Lambda\geqq a\}\to 0$. So by the same argument used before

$$|\Lambda|^{-1}\log\nu\{X_\Lambda\in[a,b)\}\to\lambda_+(a)=\sup_{a\leqq x\leqq b}\lambda_+(x).$$

Since $\lambda_+(x)<0$ for $x\geqq a$, it follows using (2.5) that $\lambda(x)=\lambda_+(x)$ on $[a,w_r]$ and hence

$$\sup_{a\leqq x\leqq b}\lambda_+(x)=\sup_{a\leqq x\leqq b}\lambda(x).$$

In case $\alpha=-$, $\gamma=+$ then by the continuity of λ_- and λ_+ on (a,b) and (2.5) it follows that there exist $c\in[a,b]$ such that $\lambda(c)=0$. Then using previous results

$$|\Lambda|^{-1}\log\nu\{X_\Lambda\in[a,b]\}\geqq|\Lambda|^{-1}\log\nu\{X_\Lambda\in[c,b]\}$$
$$\to 0=\sup_{a\leqq x\leqq b}\lambda(x).$$

c) Define $x_-=\inf\{x\in[w_1,w_r]:\lambda(x)=0\}$, $x_+=\sup\{x\in[w_1,w_r]:\lambda(x)=0\}$. We consider the case $w_1<x_-<x_+<w_r$ and leave the others to the reader. Consider partitions of $[w_1,w_r]$ into intervals $A_1=[a_0,a_1)=[w_1,a_1)$, $A_2=[a_1,a_2),\ldots,A_{M-1}=[a_{M-2},a_{M-1})$, $A_M=[a_{M-1},a_M]=[a_{M-1},w_r]$, such that for some i $a_i<x_-<a_{i+1}$ and for some j $a_j<x_+<a_{j+1}$. We suppose now that $h>0$, then

$$\max_{1\leqq k\leqq M}(\nu\{X_\Lambda\in A_k\}\exp(h|\Lambda|a_{k-1}))$$
$$\leqq E_\nu\exp(h|\Lambda|X_\Lambda)$$
$$\leqq M\cdot\max_{1\leqq k\leqq M}(\nu\{X_\Lambda\in A_k\}\exp(h|\Lambda|a_k)).$$

Hence

$$\liminf_{\Lambda\to Z^d}|\Lambda|^{-1}\log E_\nu\exp(h|\Lambda|X_\Lambda) \tag{2.6}$$
$$\geqq\max_{1\leqq k\leqq M}(h\,a_{k-1}+\liminf_{\Lambda\to Z^d}(|\Lambda|^{-1}\log\nu\{X_\Lambda\in A_k\}))$$

and

(2.7)
$$\limsup_{\Lambda\to Z^d} |\Lambda|^{-1} \log E_\nu \exp(h|\Lambda| X_\Lambda) \leqq \max_{1\leqq k\leqq M} (h a_k + \liminf_{\Lambda\to Z^d}(|\Lambda|^{-1} \log \nu\{X_\Lambda \in A_k\})).$$

The point now is that in spite of part (b) not giving us information for the terms $k=i+2, i+3, \ldots, j$, this is not important since the maxima in (2.6) and (2.7) must be achieved at some $k\in\{j+1, \ldots, M\}$. In fact the terms $k=1, \ldots, j$ are not larger than the term $k=j+1$:

i) For $i\leqq k\leqq j$

$$h a_{k-1} + \liminf_{\Lambda\to Z^d} |\Lambda|^{-1} \log \nu\{X_\Lambda\in A_k\} \leqq h a_k + \limsup_{\Lambda\to Z^d} |\Lambda|^{-1} \log \nu\{X_\Lambda \in A_k\} \leqq h a_j.$$

ii) $h a_j + \lim_{\Lambda\to Z^d} |\Lambda|^{-1} \log \nu\{X_\Lambda \in A_{j+1}\} = h a_j$.

Applying part *b* for $k=j+1, \ldots, M$ it follows that

$$\liminf_{\Lambda\to Z^d} |\Lambda|^{-1} \log E_\nu \exp\{h|\Lambda| X_\Lambda\} \geqq \max_{j+1\leqq k\leqq M} (h a_{k-1} + \lambda(a_{k-1})).$$

$$\limsup_{\Lambda\to Z^d} |\Lambda|^{-1} \log E_\nu \exp\{h|\Lambda| X_\Lambda\} \leqq \max_{j+1\leqq k\leqq M} (h a_k + \lambda(a_{k-1})).$$

Take a sequence of partitions such that $\max_{1\leqq k\leqq M} (a_k - a_{k-1}) \to 0$ to conclude the proof.

If $h<0$ the proof is analogous and if $h=0$ it is trivial.

Remark 1. The convergence of $|\Lambda|^{-1} \log E_\nu \exp(h|\Lambda| X_\Lambda)$, to a convex function $\pi(h)$ may be proven in an easier way from the FKG relations. But this approach does not give the relation between $\pi(h)$ and $\lambda(x)$ unless $\pi(h)$ is differentiable. [Si], [Pl.], [P.S.], [Ell.].

Remark 2. That the condition (1.1) may be essential can be seen with the following example: $W=\{0, 1\}$ and $\nu=(1/2)\delta_0+(1/2)\delta_1$; in this case $\lambda(x)=0$ for any $x\in[0, 1]$, but if $0<a<b<1$, $|\Lambda|^{-1} \log \nu(X_\Lambda\in[a, b]) = -\infty$.

3. Invariant Measures of Attractive Spin Systems

Proof of Theorem 2. Since the proofs of (1.9) and (1.10) are analogous we only present the first one. Divide Z^d into cubes of side N; to be precise suppose that one of these cubes is $\Gamma=\{1, \dots, N\}^d$. For $i, j \in Z^d$ write $i \sim j$ if i and j belong to the same cube in this partition. Now define a new attractive spin system with the rates

$$c_N(i, \eta)=c(i, \eta^i) \tag{3.1}$$

where η^i is the configuration defined by

$$\eta^i(j)=\begin{cases}\eta(j) & \text{if } i \sim j \\ 1 & \text{otherwise}\end{cases} \tag{3.2}$$

i.e., we consider the time evolution in Γ with $+$ boundary conditions. Denoting by $S_N(t)$ the corresponding semigroup, $\delta_1 S_N(t)$ converges weakly to a measure μ_N which has the following two nice properties:

(3.3) μ_N is stochastically larger than v_+,

(3.4) w.r.t. μ_N the spins in different cubes (of side N) are independent and moreover the distribution is the same inside each cube (μ_N is periodic).

Property (3.4) is clearly true and (3.3) follows from Corollary 1.7 of Chap. III of [Lig] by the attractiveness of $c(i, \eta)$.

If Λ is a cube of side kN, $k=1, 2, \dots$, there is a translation of it which is the union of k^d of the cubes of side N. Therefore, using (3.3), (3.4) and the large deviation theorem for independent identically distributed random variables it follows that

$$v\{X_\Lambda \geqq x\} \leqq v_+\{X_\Lambda \geqq x\} \leqq \mu_N\{X_\Lambda \geqq x\} \leqq C\, e^{-\gamma|\Lambda|},$$

provided that

$$x \geqq \rho_N = E_{\mu_N}(X_\Gamma),$$

where C and γ depend on N and x.

To complete the proof we must show that

$$\lim_{N \to \infty} \rho_N = \rho_+ . \tag{3.5}$$

For each N, let i_N be a site in $\tilde{\Gamma}=\Gamma \setminus [\sqrt{N}, N-\sqrt{N}]^d$ such that $E_{\mu_N}(i_N) \geqq E_{\mu_N}(j)$ for any $j \in \tilde{\Gamma}$. Then using (3.3)

$$\rho_+ \leqq \rho_N \leqq E_{\mu_N}(i_N) + \varepsilon_N,$$

where $\varepsilon_N \to 0$ as $N \to \infty$. For each N translate i_N to the origin and apply the same translation to Γ obtaining Γ_N. Then $\Gamma_N \to Z^d$ and by Theorem 2.7 of Chap. III of [Lig] $\lim_{N \to \infty} E_{\mu_N}(i_N) = \rho_+$.

This completes the proof (the extension to Λ whose sides are not a multiple of N is straightforward).

Generalizations

We point out now two extensions of the previous results. Instead of the mean value of the spin – X_Λ – one can consider the mean value of an increasing local (cylindrical) function f

$$Y_\Lambda(\eta)=|\Lambda|^{-1}\sum_{i\in\Lambda} f(\tau_i\,\eta)$$

where $(\tau_i\,\eta)(j)=\eta(j+i)$. The proof of Theorem 3 must be slightly modified by the inclusion of corridors between the cubes of side N. The width of these corridors can be taken as twice the range of f and w.r.t. the dynamics $S_N(t)$ the spins inside the corridors are frozen in the value $+1$.

Another extension of our methods is for occupation times. Let $(\xi_t^\nu, t\geqq 0)$ be a TIA starting from an invariant measure ν, which is FKG and non degenerate, and consider for instance the mean occupation time of the origin

$$Z_t=t^{-1}\int_0^t \xi_s^\nu(0)\,d\,s.$$

Define

$$\zeta(i)=\int_{i-1}^{i} \xi_s^\nu(0)\,d\,s.$$

$(\zeta(i), i=1, 2, \ldots)$ is FKG by Corollary 2.21 of Chap. II of [Lig], and for t integer

$$Z_t=t^{-1}\sum_{i=1}^{t}\zeta(i).$$

Theorem 1 applies with minor modification since $\zeta(i)$ assumes values in the continuum $[0, 1]$. Theorem 2 also holds – to construct μ_N reset the configuration to be identically 1 at the instants $t=N, 2N, 3N, \ldots$. The conclusion is the existence of a concave function λ: $[0, 1]\to(-\infty, 0]$ such that

a) $\lambda(x)<0$ if $x<\rho_-$ or $x>\rho_+$

b) $\lim_{t\to\infty} t^{-1}\log P(Z_t\in[a, b])=\sup_{a\leqq x\leqq b}\lambda(x)$ if $\min(\lambda(a), \lambda(b))<0$.

Examples

We consider now some examples of TIA.

1. *Basic Contact Process* [Lig].

$$c(i,\eta)=\begin{cases}1 & \text{if } \eta(i)=1\\ \beta\sum_{j:\,\|j-i\|=1}\eta(j) & \text{if } \eta(i)=0\end{cases}$$

where $\beta>0$ is *a* constant. One of the basic results for the process is that for any d, there exists a critical value $\beta_d\in(0,\infty)$ such that if $\beta<\beta_d$ the system is ergodic and the unique invariant measure is δ_0. If $\beta>\beta_d$, $\nu_-=\delta_0$ but $\nu_+\neq\delta_0$ and $0<\rho_+<1$. From Corollary 1 it follows that the "free energy" $\lambda(x)$ corresponding to ν_+ is negative for $x>\rho_+$, but this corollary gives no information for $x<\rho_+$. We prove now that for $d=1$, $\lambda(x)$ is negative for x close to 0:

$$\nu\{X_\Lambda\leqq x\}\leqq\sum\nu\{\eta(i_1)=\eta(i_2)=\ldots=\eta(i_k)=0\} \tag{3.6}$$

where k is the integer part of $(1-x)|\Lambda|$ and the sum is over the $\binom{|\Lambda|}{k}$ possible choices of k sites out of $|\Lambda|$. But it is known [D.G.], [Lig] that each term on the r.h.s. of (3.6) is bounded by $C\,e^{-\gamma k}$, where C and γ are positive. So

$$\nu\{X_\Lambda\leqq x\}\leqq\binom{|\Lambda|}{k}C\,e^{-\gamma k}.$$

Hence

$$\limsup_{\Lambda\to Z}|\Lambda|^{-1}\log\nu\{X_\Lambda\leqq x\}\leqq-(1-x)\log(1-x)-x\log x-(1-x)\gamma,$$

which converges to $-\gamma$ as $x\to0$.

We do not know whether $\lambda(x)$ is negative for every $x<\rho_+$.

2. *Basic Voter Model* [Lig].

$$c(i,\eta)=\begin{cases}(2d)^{-1}\sum\limits_{j:\,\|i-j\|=1}(1-\eta(j)) & \text{if } \eta(i)=1\\ (2d)^{-1}\sum\limits_{j:\,\|i-j\|=1}\eta(j) & \text{if } \eta(i)=0.\end{cases}$$

In $d=1,2$ there are only two extremal invariant measures: δ_0 and δ_1. In $d\geqq3$ there is a one parameter family of extremal invariant measures $\{\nu_\rho\colon 0\leqq\rho\leqq1\}$ where ρ can be chosen as the density of ν_ρ. These measures are translation invariant and FKG but their correlations decay very slowly. We will verify that for any ν_ρ with $0<\rho<1$ the corresponding "free energy" $\lambda(x)$ is identically 0. The main tool is the dual of the voter model: a system of coalescing random walks [Lig]. At each site of Λ start a simple symmetric continuous time random walk with mean holding time 1. The walkers behave independently before meeting, but when they meet they coalesce. Let N_t be the number of distinct walkers at time t; then $N_t\to N_\infty$ a.s. and by duality

$$\nu_\rho\{\eta(x)=1,\,x\in\Lambda\}=\sum_{k=1}^{|\Lambda|}\rho^k P(N_\infty=k). \tag{3.7}$$

We will use now (3.7) to show that

$$\lim_{\Lambda\to Z^d}|\Lambda|^{-1}\log\nu_\rho\{X_\Lambda=1\}=0. \tag{3.8}$$

From (3.7), for any α

$$\nu_\rho\{X_\Lambda=1\}\geqq\rho^{|\Lambda|^{\alpha/d}}P(N_\infty\leqq|\Lambda|^{\alpha/d})$$

Therefore

$$|\Lambda|^{-1}\log\nu\{X_\Lambda=1\}\geqq|\Lambda|^{(\alpha-d)/d}\log\rho+|\Lambda|^{-1}\log(1-P(N_\infty>|\Lambda|^{\alpha/d})).$$

From the techniques used in [B.C.G.1] it follows that $E(N_\infty)=O(|\Lambda|^{(d-2)/d})$; using then Chebyshev inequality

$$P(N_\infty>|\Lambda|^{\alpha/d})\leqq|\Lambda|^{-\alpha/d}E(N_\infty)=O(|\Lambda|^{(d-\alpha-2)/d}).$$

The choice $\alpha=d-1$ completes the argument for (3.8). It follows that $\lambda_+(x)=0$ for any $x\in[0,1]$. By an analogous argument $\lambda_-(x)$ is identically 0, and hence the same is true for $\lambda(x)$.

We conjecture that

$$|\Lambda|^{(-d+2)/d}\log\nu_\rho\{X_\Lambda\geqq x\}$$

converge to a non trivial limit but were not able to prove it.

3. Ergodic Systems. As already observed our results are more informative for ergodic TIAs. Various sufficient conditions for a system to be ergodic are known. For instance if the system is additive and extralineal, in the terminology of [Gri] (this means that it can be constructed with a random graph and that there are spontaneous births of particles) then it is ergodic by Theorem 2.2 of Chap. II of that book. An example in this class is a Voter model with spontaneous flips, defined by the rates:

$$c(i,\eta)=\begin{cases}(2d)^{-1}\sum_{j:\|i-j\|=1}(1-\eta(j))+\delta & \text{if } \eta(i)=1\\ (2d)^{-1}\sum_{j:\|i-j\|=1}\eta(j)+\beta & \text{if } \eta(i)=0\end{cases}$$

where β and δ are positive.

It is also known that if we add large constants to the rates of any TIA the resulting process, which is clearly a TIA, is ergodic (Theorem 4.1 of Chap. I of [Lig]).

Small Fluctuations and Large Deviations

We turn now to the relation between the large deviations and the central limit theorem. In various cases it is known for invariant measures of TIA that

$$\sigma^2=\sum_{i\in Z^d}\mathrm{Cov}_\nu(\eta(0),\eta(i))<\infty.$$

See for instance Theorem 4.22 of Chap. I of [Lig] or Theorem 2.6 of Chap. II of [Gri] (this one includes the voter model with spontaneous flips).

Assuming that ν is translation invariant and FKG, one can apply a theorem by Newman [New] to prove

$$|\Lambda|^{-1/2}\sum_{i\in\Lambda}(\eta(i)-\rho)\to \text{Normal}(0,\sigma^2)$$

in law as $\Lambda\to Z^d$, where $\rho=E_\nu(\eta(0))$. If ν is also non degenerate, $\text{Var}_\nu(\eta(0))>0$ and therefore $\sigma^2>0$. The following well known heuristic argument leads to a relation between $\lambda(x)$ and σ^2:

$$\nu\{X_\Lambda=\rho+x|\Lambda|^{-1/2}\}\sim\exp(|\Lambda|\lambda(\rho+x|\Lambda|^{-1/2}))$$
$$\sim\exp(|\Lambda|(\lambda(\rho)+\lambda'(\rho)x|\Lambda|^{-1/2}+\lambda''(\rho)x^2|\Lambda|^{-1}/2+\ldots)).$$

Assuming that $\lambda'(\rho)=0$ ($\lambda(x)$ has a maximum at ρ) and remembering that $\lambda(\rho)=0$, it follows

$$\nu\{X_\Lambda=\rho+x|\Lambda|^{-1/2}\}\sim\exp(\lambda''(\rho)x^2/2).$$

Hence

$$\sigma^2=-(\lambda''(\rho))^{-1}. \tag{3.9}$$

This relation is indeed true when ν is Bernoulli (independent case) and for Ising models at high temperature or with non zero external field (here ν is the unique Gibbs state) [Ell.; Leb1], but in general it is not even true that $\lambda''(\rho)$ exists. For instance consider the nearest neighbor ferromagnetic Ising model in two or more dimensions at low temperature and without external field; the corresponding reversible Glauber dynamics is a TIA (on $\{-1,1\}^{Z^d}$ instead of $\{0,1\}^{Z^d}$) and the Gibbs measures are invariant for this dynamics). In this case ν_- and ν_+ are different but have the same function $\lambda(x)$, which is null on $[\rho_-,\rho_+]$ and negative outside this interval. With respect to the measure ν_+ the correlations decay exponentially, so that $\sigma^2\in(0,\infty)$. But it is clear that the second derivative of $\lambda(x)$ at $x=\rho_+$ coming from the left is 0. It turns out that coming from the right this second derivative is σ^2 [Leb1] (σ^2 is the susceptibility).

An even more interesting example is the case of a spin 1 Ising model [B.K.L.; Sl.] on Z^d, $d\geqq 2$, with energy

$$U(\eta)=\tfrac{1}{2}\sum_{\|i-j\|=1}(\eta(i)-\eta(j))^2+g\sum_i(\eta(i))^2,$$

$\eta(i)=-1$, 0, 1, $g\in R$ is a parameter. If the temperature is low enough there is a value of g such that there are three extremal Gibbs states: μ_+, μ_- and μ_0 with respective densities, $\rho_+=-\rho_->0$, and $\rho_0=0$. These measures are translation invariant, FKG and have exponentially decaying correlations. Since they are Gibbs measures corresponding to the same interaction they have the same function $\lambda(x)$, which is null for $\rho_-\leqq x\leqq\rho_+$ and negative outside this interval. So $\lambda(x)$ has derivatives of all orders equal to zero at $x=\rho_0=0$. But if the temperature is positive μ_0 is not degenerate and therefore $0<\sigma^2=\sum_{i\in Z^2}\text{Cov}_{\mu_0}(\eta(0),\eta(i))<\infty$. Similar situations occur for higher states Potts models.

We leave as an open problem the question whether (3.9) holds for the unique invariant measure of ergodic TIA. One should note that at the critical temperature of an Ising system $\sigma^2=\infty$, but then $\lambda''(\rho)=0$, so that (3.9) holds.

4. Perturbation by an External Field

Proof of Theorem 3. Define

$$\Pi_{\Lambda,h}(h')=|\Lambda|^{-1}\log E_{\nu_{\Lambda,h}}(\exp(h'|\Lambda|X_\Lambda))$$

$$\Pi_\Lambda(h')=\Pi_{\Lambda,0}(h')=|\Lambda|^{-1}\log Z(\Lambda,h')$$

Then

$$\Pi_{\Lambda,h}(h')=|\Lambda|^{-1}\log\sum_{\eta\in w^\Lambda}(Z(\Lambda,h))^{-1}\nu_\Lambda(\eta)\exp((h+h')|\Lambda|X_\Lambda)$$

$$=\Pi_\Lambda(h+h')-\Pi_\Lambda(h).$$

By the argument given in part c of Theorem 1, $\Pi_\Lambda(h)\to\Pi(h)$ as $|\Lambda|\to\infty$. So

$$\lim_{\Lambda\to Z^d}\Pi_{\Lambda,h}(h')=\Pi(h+h')-\Pi(h).$$

Hence

$$\inf_{h'}(\Pi_h(h')-h'x)=xh-\Pi(h)$$
$$+\inf_{h'}(\Pi(h+h')-x(h+h'))=\lambda(x)+xh-\Pi(h)=\lambda_h(x)$$

and by a standard argument using a Chebyshev inequality (see [Ell], Theorem II.6.1(b)),

$$\limsup_{\Lambda\to Z^d}|\Lambda|^{-1}\log\nu_{\Lambda,h}\{X_\Lambda\in[a,b]\}\leqq\sup_{a\leqq x\leqq b}\lambda_h(x).$$

For the lower bound consider first $h>0$. Fix a point $x\in(a,b)$, then for $\varepsilon>0$ small enough

$$\liminf_{\Lambda\to Z^d}|\Lambda|^{-1}\log\nu_{\Lambda,h}\{X_\Lambda\in[a,b]\}$$
$$\geqq\liminf_{\Lambda\to Z^d}|\Lambda|^{-1}\log\nu_{\Lambda,h}\{X_\Lambda\in[x-\varepsilon,x+\varepsilon]\}$$
$$\geqq\liminf_{\Lambda\to Z^d}|\Lambda|^{-1}\log[(Z(\Lambda,h))^{-1}\exp(|\Lambda|h(x-\varepsilon))\,\nu\{X_\Lambda\in[x-\varepsilon,x+\varepsilon]\}]$$
$$=-\Pi(h)+h(x-\varepsilon)+\sup_{x-\varepsilon\leqq y\leqq x+\varepsilon}\lambda(y).$$

Making $\varepsilon\to 0$ the r.h.s. above converges to $\lambda_h(x)$. Since x is arbitrary

$$\liminf_{\Lambda\to Z^d}|\Lambda|^{-1}\log\nu_{\Lambda,h}\{X_\Lambda\in[a,b]\}\geqq\sup_{a<x<b}\lambda_h(x)=\sup_{a\leqq x\leqq b}\lambda_h(x).$$

The case $h<0$ is analogous.

5. Discussion

Various approaches are known for proving large deviations theorems for dependent random variables. It is natural to ask (as we did) why we could not simply verify the hypothesis of some of these general theorems for the cases considered here.

Lanford [Lan] proved that (1.2) holds for any $w_1 < a < b < w_r$ for Gibbs measures. So if we were able to prove that ν is Gibbs w.r.t. some nice enough potential we would be finished. Unfortunately the only condition that we know to prove that a measure is Gibbs – the continuity of some conditional expectations (see [Sul]) – seems hard to verify in our cases.

Many approaches do not work for our problem because we consider random variables $\eta(i)$ indexed by Z^d while they consider random variables indexed by Z or R. That is the case of approaches by Accardi and Olla [A.O.], Olla [Ol.] and Orey [Or.]. Even in the case $d=1$ their conditions seem to be difficult to verify in our case and to be related to Sullivan's condition [Sul.] mentioned before.

Finally the approach due to Sievers, Plachky, Steinbach and Ellis [Si; Pl.; P.S.; Ell.] and used in other applications to infinite particle systems [C.G.1; C.G.2; B.C.G.2] requires that one first proves the convergence of $|\Lambda|^{-1} \log E_\nu(\exp(h|\Lambda| X_\Lambda))$ to a convex function $\Pi(h)$ which is *differentiable*. The problem is that we do not know how to prove this last condition. We should remark that unfortunately the method that we used is not suitable for proving large deviations theorems with "fat" tails like in [C.G.1; C.G.2; B.C.G.2] and as we expect for the voter model.

Acknowledgements. This paper benefited from discussions with many colleagues. The approach to the proof of Theorem 2 emerged from conversations with Rick Durrett. The spin 1 Ising model as an example for which (3.9) does not hold was suggested to us by Michael Aizenman. We thank also T. Cox, S. Goldstein, H. Kesten, T. Liggett, S. Olla and D. Ruelle for useful discussions. R.H.S. thanks also the warm hospitality of the Rutgers Mathematics Department.

References

[A.O.] Accardi, L., Olla, S.: Donsker and Varadhan's theory for stationary processes. Preprint

[B.C.G.1] Bramson, M., Cox, T., Griffeath, D.: Consolidation rates for two interacting systems in the plane. Preprint

[B.C.G.2] Bramson, M., Cox, T., Griffeath, D.: Occupation time large deviations of the voter model. Preprint

[B.K.L.] Bricmont, J., Kuruda, K., Lebowitz, J.L.: First order phase transitions in lattice and continuous systems: extension of Pirogov-Sinai theory. Commun. Math. Phys. **101**, 501–538 (1985)

[C.G.1] Cox, T., Griffeath, D.: Large deviations for Poisson systems of independent random walks. Z. Wahrscheinlichkeitstheor. Verw. Geb. **66**, 543–558 (1984)

[C.G.2] Cox, T., Griffeath, D.: Occupation times for critical branching Brownian motions. Ann. Probab. **13**, 1108–1132 (1985)

[D.G.] Durrett, R., Griffeath, D.: Supercritical contact processes on Z. Ann. Probab. **11**, 1–15 (1983)

[Ell.] Ellis, R.: Entropy, large deviations and statistical mechanics. Berlin Heidelberg New York: Springer 1985

[F.K.G.] Fortuin, C.M., Kasteleyn, P.W., Ginibre, J.: Correlation inequalities on some partially ordered sets. Commun. Math. Phys. **22**, 89–103 (1971)

[G.K.I.] Goldstein, S., Kipnis, C., Ianiro, N.: Stationary states for mechanical systems with stochastic boundary conditions. J. Stat. Phys. **41**, 915–939 (1985)

[G.L.P.] Goldstein, S., Lebowitz, J.L., Presutti, E.: Mechanical systems with stochastic boundaries. In: Fritz, J., Lebowitz, J.L., Szasz, D. (eds.) Random Fields. Amsterdam: North-Holland 1981

[Gra] Gray, L.: The critical behavior of a class of simple interacting systems – a few answers and a lot of questions. In: Durrett, R. (ed.) Contemporary mathematics, vol. 41, pp. 149–160. Providence: Am. Math. Soc. 1985

[Gri] Griffeath, D.: Additive and cancelative interacting particle systems. Lect. Notes Math., vol. 724. Berlin Heidelberg New York: Springer 1979

[K.L.S.] Katz, S., Lebowitz, J.L., Spohn, H.: Nonequilibrium steady states of stochastic lattice gas models of fast ionic conductors. J. Stat. Phys. **34**, 497–537 (1984)

[Kün] Künsch, H.: Non reversible stationary measures for infinite interacting particle systems. Z. Wahrscheinlichkeitstheor. Verw. Geb. **66**, 407–424 (1984)

[Lan.] Lanford, O.E.: Entropy and equilibrium states in classical statistical mechanics. Lect. Notes Phys., vol. **20**, pp. 1–113. Berlin Heidelberg New York: Springer 1971

[Leb1] Lebowitz, J.L.: Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems. Commun. Math. Phys. **28**, 313–321 (1972)

[Leb2] Lebowitz, J.L.: Exact results in nonequilibrium statistical mechanics: Where do we stand? Prog. Theor. Phys. **64**, 35–49 (1978)

[Lig] Liggett, T.M.: Interacting particle systems. Berlin Heidelberg New York: Springer 1985

[New] Newman, C.M.: Normal fluctuations and the FKG inequalities. Commun. Math. Phys. **74**, 119–128 (1980)

[Ol.] Olla, S.: Large deviations for almost Markovian processes. Probab. Th. Rel. Fields **16**, 395–409

[Or.] Orey, S.: Large deviations and Shanon-McMillan theorems. Seminar on Stochastic Processes, 1984. Boston: Birkäuser 1986

[Pl.] Plachky, D.: On a theorem by G.L. Sievers. Ann. Math. Statist. **42**, 1442–1443 (1971)

[P.S.] Plachky, D., Steinbach, J.: A theorem about probabilities of large deviations with an application to queuing theory. Per. Math. Hungarica **6**, 343–345 (1975)

[Ru1] Ruelle, D.: Statistical mechanics. Rigorous results. Menlo Park: Benjamin 1969

[Ru2] Ruelle, D.: Thermodynamic formalism. Encyclopedia of Mathematics and Its Applications, vol. 5. Reading, Mass.: Addison-Wesley 1978

[Si] Sievers, G.L.: On the probability of large deviations and exact slopes. Ann. Math. Statist. **40**, 1908–1921 (1969)

[Sin] Sinai, Ya.G.: Theory of phase transitions: Rigorous results. Oxford New York: Pergamon Press 1982

[Sl.] Slawny, J.: Low temperature properties of classical lattice systems: phase transitions and phase diagrams. In: Domb, C., Lebowitz, J.L. (eds.) Phase Transitions and Critical Phenomena, Vol. 11. Amsterdam: North Holland 1987

[Sul.] Sullivan, W.G.: Potentials for almost Markovian random fields. Commun. Math. Phys. **33**, 61–74 (1973)

Received April 24, 1986

Note added in proof. 1. After this paper was finished, we learned from "Grimmet, G.: Large deviations in subadditive processes and first-passage percolation. In: Durret R. (ed.) Particle Systems, Random Media, and Large Deviations, pp. 175–194, Providence: American Mathematical Society 1984" that Hammersley and Kingsman worked already on large deviation properties under a condition similar (but not equivalent) to FKG (they considered so-called superconductive processes). Grimmett also remarks in the paper that the techniques used in these works can be adapted to the case of FKG sequences of random variables to prove part of our Theorem 1.

2. After this paper was finished, R. Durrett and one of us (R.H.S.) proved for the upper invariant measure of the contact process the function $\lambda(x)$ is negative for $x<\rho_+$. This result is contained in the paper "Large deviations for the contact process and two dimensional percolation", which is to appear in this journal.