Rendiconti del Seminario Matematico della Università di Padova

Paolo Dai Pra

A minimum entropy problem for stationary reversible stochastic spin systems on the infinite lattice

Rendiconti del Seminario Matematico della Università di Padova, tome 92 (1994), p. 179-194

<http://www.numdam.org/item?id=RSMUP_1994__92__179_0>

REND. SEM. MAT. UNIV. PADOVA, Vol. 92 (1994)

A Minimum Entropy Problem for Stationary Reversible Stochastic Spin Systems on the Infinite Lattice.

PAOLO DAI PRA (*)

1. Introduction.

In this paper we formulate and solve a variational principle that belongs to a wide and very much studied class of problems that, throughout this work, will be called *minimum entropy problems for stochastic processes*. We start by giving a basic example of a problem in such class.

Suppose we are given a stochastic process $X = \{X_t : t \in [0, 1]\}$ taking values in a measurable space E. The path space $\mathcal{D} = E^{[0, 1]}$ is provided with the natural σ-field generated by cylinder sets. The process X induces on $\mathcal{D}$ a probability measure that we denote by P. In the rest of the paper we often identify E-valued stochastic processes and probability measures on $\mathcal{D}$. The collection of probability measures on $\mathcal{D}$ is denoted by $\mathcal{M}(\mathcal{D})$.

Now let Q be another element of $\mathcal{M}(\mathcal{D})$. The *relative entropy* of Q with respect to P is defined to be

$$h(Q|P) = \int \left(\log \frac{dQ}{dP}\right) dQ = E^Q \left\{\log \frac{dQ}{dP}\right\}, \tag{1}$$

where $h(Q|P)$ is assumed to be $+\infty$ if $Q \not\ll P$ or $dQ/dP \notin L^1(Q)$. Notice that, by Jensen's inequality, $h(Q|P) \geqslant 0$ and $h(Q|P) = 0$ if and only if $Q = P$. Now let A be a subset of $\mathcal{M}(\mathcal{D})$, that will be called the *constraint set*. A minimum entropy problem consists in minimizing $h(Q|P)$ under the constraint set $Q \in A$ and, if possible, in finding a $Q \in A$ that realizes the minimum.

(*) Indirizzo dell'A: Università di Padova, Dipartimento di Matematica, Via Belzoni 7, 35100 Padova, Italy and Rutgers University, Departement of Mathematics, New Brunswick, NJ 08903, U.S.A.

A minimum entropy problem is therefore identified once a stochastic process and a constraint set are given. The type of constraint set which is most often studied in the literature is the one that gives rise to the so called Schroedinger problem ([1,5,7,8,10,11], and it consists in fixing the initial and final marginal of Q. To be more precise, let us denote by $\pi_t Q$, $t \in [0, 1]$, the projection of Q at time t, i.e. for $G \in E$ measurable

$$\pi_t Q(G) = Q\{x \in \mathcal{D}: \ x_t \in G\}.$$

Moreover let μ, ν be two probability measures on E. In the Schroedinger problem the constraint set is of the form

$$A = \{Q \in \mathcal{M}(\mathcal{D}): \ \pi_0 Q = \mu, \ \pi_1 Q = \nu\}.$$

Roughly speaking, to solve the Schroedinger problem means to construct a stochastic «bridge» from μ to ν which is «as close as possible» to the given process P. The «closeness» we are talking about in this statement is well explained by the *large deviation* interpretation of an entropy problem [5], that we now briefly sketch.

Consider the product space $\Omega \equiv \mathcal{D}^{\boldsymbol{N}}$, and suppose we provide it with the product measure $\boldsymbol{P} \equiv \otimes_{\boldsymbol{N}} P$. In other words we are given a collection $X(i) = \{X_t(i): \ t \in [0, 1]\}$, $i \in \boldsymbol{N}$, of independent copies of X. For $n \in \boldsymbol{N}$ and $\omega \in \Omega$ we define the empirical measure $L_{n,\omega} \in \mathcal{M}(\mathcal{D})$ by

$$L_{n,\omega} = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\omega(i)} \tag{2}$$

where $\delta_{\omega(i)}$ is the Dirac measure concentrated at $\omega(i) \in \mathcal{D}$. By the Ergodic Theorem

$$\lim_{n \to \infty} L_{n,\omega} = P \tag{3}$$

for $\boldsymbol{P}$-almost every ω, where the limit is taken with respect to some suitable topology on $\mathcal{M}(\mathcal{D})$ (for instance the one induced by duality by the bounded measurable functions on $\mathcal{D}$). Now let $C \subset \mathcal{M}(\mathcal{D})$ be such that $P \notin C$. A refinement of (3) is given by the following large deviation principle, which holds under some regularity assumptions on C:

$$\lim_{n \to \infty} \frac{1}{n} \log \boldsymbol{P}\{\omega: \ L_{n,\omega} \in C\} = - \inf_{Q \in C} h(Q|P).$$

In particular, for $Q \in \mathcal{M}(\mathcal{D})$ and N_Q a «small» set containing Q (small with respect to the variations of $h(\cdot|P)$) one has, roughly,

$$\boldsymbol{P}\{\omega: \ L_{n,\omega} \in N_Q\} \approx e^{-nh(Q|P)}. \tag{4}$$

The asymptotic estimate (4) yields the following interpretation of the solution of an entropy problem with constraint A: the solution Q of the problem is, among the elements of A, the evolution which is the most likely to be «observed», in the sense that, in the limit $n \to \infty$, it maximizes the probability that $L_{n,\omega}$ is close to Q.

The entropy problems we have described so far make sense for an arbitrary process X. The picture is somewhat different when we know *a priori* that the process $X = \{X(t)\colon t \in \boldsymbol{R}\}$, now defined on the whole real line, is stationary, i.e. the probability measure P it induces on the path space $\mathcal{D} = E^{\boldsymbol{R}}$ is invariant under the maps $\{\theta_t\colon t \in \boldsymbol{R}\}$ where, for $x \in \mathcal{D}$,

$$(\theta_t x)_s = x_{t+s}\,.$$

The stationarity of the process suggests a different way of computing its statistics. If we are given an arbitrary process, its statistics can be computed by producing several independent copies of the process and then by evaluating the empirical averages defined in (2). For ergodic stationary processes this is not necessary: by letting the process run for a long enough time, all the statistics can be determined. That is to say that, by defining

(5) $$R_{n,x} = \frac{1}{n}\int_0^n \delta_{\theta_t x}\,dt \in \mathcal{M}(\mathcal{D}),$$

we have

(6) $$\lim_{n\to\infty} R_{n,x} = P$$

for P-a.e. $x \in \mathcal{D}$. Under some assumptions on P, the exponential rate of convergence in (6) is controlled by a suitably defined entropy function, $H(Q\,|\,P)$, i.e., for $A \subset \mathcal{M}(\mathcal{D})$,

(7) $$P\{x\colon R_{n,x} \in A\} \approx e^{-n \inf_{Q\in A} H(Q|P)}\,.$$

We do not specify here the conditions on P that are usually assumed, both because they are not relevant for the specific model we are going to consider, and because (7) is believed to be true in much greater generality. The function $H(Q\,|\,P)$ is defined as follows. Let $\mathcal{F}^-$ be the σ-field in $\mathcal{D}$ generated by $\{x_t\colon t \leq 0\}$, and denote by Q_ω (resp. P_ω) the regular conditional probability distribution (r.c.p.d.) of Q (resp. P) with respect to $\mathcal{F}^-$. Moreover we let $\mathcal{F}_1$ to be the σ-field generated by $\{x_t\colon 0 \leq$

$\leq t \leq 1\}$. Then we define

$$H(Q|P) = \int \log\left(\frac{dQ_\omega}{dP_\omega} \,|\, \mathcal{F}_1\right) Q(d\omega)$$

for $Q \in \mathcal{M}(\mathcal{D})$ stationary, $Q_{\omega|\mathcal{F}_1} \ll P_{\omega|\mathcal{F}_1}$ and $\log\left(\frac{dP_\omega}{dQ_\omega} \,|\, \mathcal{F}_1\right) \in L^1(Q)$, and $H(Q|P) = +\infty$ otherwise.

As we did for (4), the large deviation estimate (7) gives rise to another class of entropy problems, where the entropy function is now H instead of h. Thus, given a constraint $A \subset \mathcal{M}(\mathcal{D})$, we want to minimize $H(Q|P)$ for Q varying in A. A minimizer Q^* will be, among the elements of A, the one which is more likely to be close to $R_{n,\omega}$, in the limit $n \to \infty$. The type of constraint we will be interested in is a very natural one for stationary processes, and consists in fixing the one-time distribution. In other words, for μ a probability measure on E, we let

$$A = \{Q \in \mathcal{M}(\mathcal{D}): \pi_t Q = \mu \;\; \forall t \in \boldsymbol{R}\}.$$

The minimum function $I(\mu|P) \equiv \inf_{Q \in A} H(Q|P)$ has a further interesting interpretation. Let us define the *empirical measure*

$$L_{n,x} = \frac{1}{n}\int_0^n \delta_{x_t}\, dt \in \mathcal{M}(E)$$

where $\mathcal{M}(E)$ is the set of probability measures on E. Then it can be derived by (7) that, for $D \subset \mathcal{M}(E)$ sufficiently regular,

$$P\{x: L_{n,x} \in D\} \approx e^{-n \inf_{\mu \in D} I(\mu|P)}.$$

In other words $I(\mu|P)$ is the rate function that controls the large deviations of the empirical measure $L_{n,x}$.

The purpose of this paper is to solve a minimum entropy problem for a class of stationary Markov processes which is of relevance in Statistical Mechanics. As for the examples we have discussed before, the entropy function we will introduce comes from a large deviation problem, that will be stated in Section 3. The main feature of our model is the infinite dimensionality of the state space; for this reason when we define the empirical averages, we do not only average over the time but also over the different components of the process. This will lead to the definition of an entropy function $H(Q|P)$ that controls the large deviations of *space-time* empirical averages. We then ask for the minimum of $H(Q|P)$ under the constraint $\pi_t Q = \mu$, $\forall t \in \boldsymbol{R}$. Under some assumptions on the probability measure μ we will be able to find the minimizer, by showing that it is Markovian and giving its Markov generator explicitely.

In Section 2 we introduce our model, and in Section 3 we summarize the known large deviations results. The related minimum entropy problem is stated and solved in Sections 4 and 5.

2. Stochastic spin systems.

In this section we introduce the class of stochastic processes we will be dealing with in this paper. We let $X = \{-1, +1\}$ to be the set of spin values. The Markov processes we are going to define take value on $X^{\mathbf{Z}^d}$, i.e. for any site i of the d-dimensional lattice $\mathbf{Z}^d$ there is an associated spin value. The updating mechanism is specified by assigning a nonnegative function $c(i, \sigma)$, defined for $i \in \mathbf{Z}^d$, $\sigma \in X^{\mathbf{Z}^d}$. The probability of changing the sign to the spin at the site i during a time interval of length Δt, conditioned to the knowledge of the whole configuration at time t, is given by

$$P\{\sigma_{t+\Delta t}(i) = -\sigma_t(i) \mid \sigma_t\} = c(i, \sigma_t)\Delta t + o(\Delta t).$$

Moreover spins at different sites are updated independently. In particular the probability of changing the spin at two different sites in the same interval $[t, t+\Delta t]$ is $o(\Delta t)$. The functions $\{c(i, \cdot): i \in \mathbf{Z}^d\}$ are usually called *flip rates*.

The informal definition we have just given can be made rigorous as follows. First of all we provide $X^{\mathbf{Z}^d}$ with the product of the discrete topology on X. The corresponding space of real continuous functions is denoted by $\mathcal{C}(X^{\mathbf{Z}^d})$; it becomes a Banach space with the usual sup-norm. We say that a function f: $X^{\mathbf{Z}^d} \to \mathbf{R}$ is local if its dependence on $\sigma \in X^{\mathbf{Z}^d}$ is only through $\{\sigma(i): i \in \Lambda\}$, where Λ is some finite subset of $\mathbf{Z}^d$. We denote by $\mathcal{D}$ the set of local functions. On $\mathcal{D}$ we can define the following operator:

$$L^c f(\sigma) = \sum_{i \in \mathbf{Z}^d} c(i, \sigma)[f(\sigma^i) - f(\sigma)]$$

where

$$\sigma^i(j) = (-1)^{\delta_{i,j}} \sigma(j).$$

It is proved in [6] that, under the assumption

$$\sup_{i \in \mathbf{Z}^d} \sum_{j \in \mathbf{Z}^d} \sup_{\eta \in X^{\mathbf{Z}^d}} |c(i, \eta) - c(i, \eta^j)| < \infty, \tag{8}$$

the closure of L^c in $\mathcal{C}(X^{Z^d})$ generates a Markov semigroup. Moreover the corresponding Markov process is a Feller process (i.e. the elements U_t, $t > 0$ of the semigroup map bounded measurable functions into continuous functions). Notice that condition (8) essentially says that $c(i, \sigma)$ does not depend too much on the spin of sites that are far from i. We notice that (8) is satisfied when the flip rates are *translation invariant* (i.e. $c(i, \sigma) = c(0, \theta_i \sigma)$, where $\theta_i \sigma(j) = \sigma(i+j)$) and *local* (i.e. $c(0, \sigma)$ depends only on $\{\sigma(i): i \in \Lambda\}$ where Λ is a finite subset of $\boldsymbol{Z}^d$). Only these type of models will be considered in the rest of the paper.

3. Large deviations.

In this section some of the large deviations results obtained in [2,3] are summarized. We assume $c: \{-1, 1\}^{Z^d} \to \boldsymbol{R}^+$ is a local and strictly positive function. As we have seen in Section 2 the operator

$$L^c f(\sigma) = \sum_{i \in Z^d} c(\theta_i \sigma)[f(\sigma^i) - f(\sigma)]$$

is the generator of a Feller semigroup. We denote by $\{P^c_{0,\xi}: \xi \in \{-1, 1\}^{Z^d}\}$ the corresponding family of conditional probability measures. In particular, for $c \equiv 1$, we write $P_{0,\xi}$ in place of $P^1_{0,\xi}$. Notice that $P_{0,\xi}$ is simply the product measure $\prod_{i \in Z^d} P_{0,\xi(i)}$, $P_{0,\xi(i)}$ being the Markov family of a Poisson-spin process with intensity one. For obvious reasons the process generated by L^c with $c \equiv 1$ is called *non-interacting* spin system. It is well known that the trajectories of such processes are, with probability one, elements of $\Omega = D(\boldsymbol{R}, \{-1, 1\}^{Z^d})$, the set of right continuous with left limit functions from $\boldsymbol{R}$ to $\{-1, 1\}^{Z^d}$, where the topology on $\{-1, 1\}^{Z^d}$ is the product of the discrete topology. We also provide Ω with the Skorohod topology (see [4]) and the corresponding Borel σ-field, where the measures we will be considering are defined. On Ω we define the family of space-time shift maps $\{\theta_{t,n}: t \in \boldsymbol{R}, n \in \boldsymbol{Z}^d\}$ defined by

$$(\theta_{t,n}\omega)_s(i) = \omega_{s+t}(i+n).$$

DEFINITION 3.1. *A probability measure Q on Ω is said to be stationary if it is invariant for all the maps $\theta_{t,n}$.*

We denote by $\mathcal{M}_s(\Omega)$ the set of stationary measures, provided with the weak topology. The Borel sets for this topology provide Ω with a structure of measurable space.

In what follows we let $V_n = \{i \in \mathbf{Z}^d : i_j = 0, 1, \ldots, n-1, \forall j = 1, 2, \ldots, d\}$. Given $\omega \in \Omega$ we define its *n-periodic version* ω^n as follows:

$$\omega_t^n(i) = \omega_t(i) \quad \text{for } 0 \leq t \leq n, \quad i \in V_n$$

$$\omega_{t+hn}^n(i + kn) = \omega_t^n(i) \quad \text{for } h \in \mathbf{Z}, \quad k \in \mathbf{Z}^d$$

where

$$kn = (k_1 n, \ldots, k_d n).$$

For $\Gamma \subset \mathbf{R} \times \mathbf{Z}^d$ we let $\mathcal{F}_\Gamma$ to be the σ-field of subsets of Ω generated by the projections $\{\pi_{t,i} : (t, i) \in \Gamma\}$, where $\pi_{t,i}(\omega) = \omega_t(i)$. Sometimes we will use, for $\mathcal{F}_\Gamma$, the notation $\sigma\{w_t(i) : (t, i) \in \Gamma\}$. In the following definition we denote by $\mathcal{B}(\Omega)$ the set of bounded measurable functions $\Omega \to \mathbf{R}$.

DEFINITION 3.2. *Let $\omega \in \Omega$ and $\phi \in \mathcal{B}(\Omega)$. The n^{th} empirical process $R_{n,\omega}$ is the element of $\mathcal{M}_s(\Omega)$ whose expectations are defined as follows:*

$$E^{R_{n,\omega}}(\phi) = \frac{1}{n^{d+1}} \sum_{i \in V_n} \int_0^n \phi(\theta_{s,i}\omega^n)\, ds. \tag{9}$$

Notice that, in order to make $R_{n,\omega}$ stationary, it is essential to use the n^{th}-periodic version of ω in (9). We also remark that the map

$$\Omega \to \mathcal{M}_s(\Omega),$$

$$\omega \mapsto R_{n,\omega},$$

is $\mathcal{F}_{[0,n] \times V_n}$-measurable. In the rest of the paper the σ-field $\mathcal{F}_{[0,n] \times V_n}$ will be simply denoted by $\mathcal{F}_n$.

Some more notations are now needed. We introduce on $\mathbf{Z}^d$ the lexicographic total order, and denote by $<$ the corresponding order relation. Consider the set

$$\Gamma^- = \{(t, i) \in \mathbf{R} \times \mathbf{Z}^d : t \leq 1, i < 0 \text{ or } t \leq 0, i \in \mathbf{Z}^d\}.$$

For $Q \in \mathcal{M}_s(\Omega)$ we let Q_ω to denote the regular conditional probability distribution (r.c.p.d.) of Q with respect to $\mathcal{F}^-$. In the following definition $(dQ_\omega / dP_{0,\omega_0})|_{\mathcal{F}_1}$ denotes the Radon-Nikodym derivative of the indicated measures restricted to the σ-field $\mathcal{F}_1$.

DEFINITION. *Let $Q \in \mathcal{M}_s(\Omega)$. The relative entropy of Q with respect* *to the Markov family* $\{P_{0,\xi}: \xi \in X^{\mathbf{Z}^d}\}$ *is defined by*

$$H(Q) = E^Q\left\{\log\left(\frac{dQ_\omega}{dP_{0,\omega_0}} \mid \mathcal{F}_1\right)\right\} \tag{10}$$

where $H(Q) = +\infty$ *if the Radon-Nikodym derivative in* (10) *is not defined or its logarithm is not in* $L^1(Q)$.

In what follows, for $\omega \in \Omega$, $i \in \mathbf{Z}^d$, $t \in \mathbf{R}$, we let

$$N_t(i) = \sum_{0 \leq s \leq t} \frac{|\omega_{t^-}(i) - \omega_t(i)|}{2} + \delta_{1,\omega_0(i)}\,.$$

Notice that $\omega_t(i) = (-1)^{N_t(i)}$. The main result of [2] is the following Large Deviation Principle.

THEOREM 1. *Let A be a Borel measurable subset of $\mathcal{M}_s(\Omega)$, and denote with $\mathring{A}$ and $\bar{A}$ its interior and its closure respectively. Then, for every* $\xi \in X^{\mathbf{Z}^d}$

$$-\inf_{Q \in \mathring{A}} H^c(Q) \leq \liminf_{n \to \infty} \frac{1}{n^{d+1}} \log P^c_{0,\xi}\{R_{n,\omega} \in A\} \leq$$

$$\leq \limsup_{n \to \infty} \frac{1}{n^{d+1}} \log P^c_{0,\xi}\{R_{n,\omega} \in A\} \leq -\inf_{Q \in \bar{A}} H^c(Q)$$

where

$$H^c(Q) = H(Q) - E^Q\left[1 - c(\omega_0) - \int_0^1 \log c(\omega_{t^-})dN_t(0)\right].$$

In a large deviation principle it is particularly relevant to determine the zeroes of the rate function $H^c(\cdot)$. The following is the main result contained in [3]. We need to use the σ-field $\mathcal{F}^p = \sigma\{\omega_t(i): t \leq 0\}$ and, for a given $Q \in \mathcal{M}_s(\Omega)$, we let Q^p_ω denote its r.c.p.d. with respect to $\mathcal{F}^p$.

THEOREM 2. *For every $Q \in \mathcal{M}_s(\Omega)$ we have $H^c(Q) \geq 0$ and $H^c(Q) = 0$ if and only if $Q^p_\omega = P^c_{0,\omega_0}$, Q-a.s. . In other words $H^c(Q) = 0$ if and only if Q is a stationary Markovian measure generated by L^c.*

The following proposition, also proved in [3], establishes some properties of the rate function $H^c(\cdot)$ that will be used later in this paper.

THEOREM 3. *The rate function H^c: $\mathcal{M}_s(\Omega) \to \mathbf{R}^+$ is lower semicontinuous and has compact level sets, i.e. for every $l > 0$ the set $\{Q \in \mathcal{M}_s(\Omega): H^c(Q) \leq l\}$ is compact in the weak topology. Moreover*

$$H^c(Q) = \lim_{n \to \infty} \frac{1}{n^{d+1}} E^Q \left\{ \log \left(\frac{dQ_\omega^p}{dP_{0,\omega_0}^c} \mid \mathcal{F}_n \right) \right\}.$$

4. The I function and its basic properties.

Now we are ready to state the minimum entropy problem we want to solve. Let μ be an element of $\mathcal{M}_s(X^{\mathbf{Z}^d})$, i.e. a probability measure on $X^{\mathbf{Z}^d}$ which is invariant under the shift maps θ_i, $i \in \mathbf{Z}^d$. Our goal is to minimize the function $H^c(Q)$ under the constraint $\pi_t Q = \mu$ for every $t \in \mathbf{R}$. In this section we give the basic properties and the large deviation interpretation of the minimum function $I(\mu) \equiv \inf\{H^c(Q): \pi_t Q = = \mu\}$. In the next section, under some assumption on the model and on μ, we will be able to actually compute $I(\mu)$, and a minimizer Q^* (i.e. $\pi_t Q^* = \mu$ and $H^c(Q^*) = I(\mu)$).

The large deviation theory we have summarized in the previous section allows to determine the asymptotic behavior of quantities of the form

$$P_{0,\xi}^c\{R_{n,\omega}(F) \in A\} \tag{11}$$

for any bounded measurable, $\mathbf{R}^m$-valued function F on the path space Ω, and any Borel set $A \subset \mathbf{R}^m$, $m > 0$. A function F of the type described above is often called an *observable*.

A class of observable that play a significant role are those for which there exists a function f: $X^{\mathbf{Z}^d} \to \mathbf{R}^m$ such that

$$F(\omega) = f(\omega_0). \tag{12}$$

In other words F depends only on the value of the trajectory ω at a given time. Notice that, for any bounded measurable f: $X^{\mathbf{Z}^d} \to \Omega$ we can define

$$L_{n,\omega}(f) = R_{n,\omega}(F) \tag{13}$$

where F is defined by (12). It is easy to check that, for $n \in \mathbf{N}$, $\omega \in \Omega$, $L_{n,\omega}$ can be identified with a measure on $X^{\mathbf{Z}^d}$ which is invariant under the shift maps $\{\theta_i: i \in \mathbf{Z}\}$, i.e. $L_{n,\omega} \in \mathcal{M}_s(X^{\mathbf{Z}^d})$. Therefore, after having provided $\mathcal{M}_s(X^{\mathbf{Z}^d})$ with the weak topology one can define a family of measures $\{\Gamma^c_{n,\xi}: n \in \mathbf{N}, \xi \in X^{\mathbf{Z}^d}\}$ on $\mathcal{M}_s(X^{\mathbf{Z}^d})$ by

$$\Gamma^c_{n,\xi}(A) = P^c_{0,\xi}\{\omega: L_{n,\omega} \in A\} \tag{14}$$

and ask about the large deviation properties of $\Gamma^c_{n,\xi}$. The first fact we state is an immediate consequence of the *contraction principle* (see [9]).

THEOREM 4. *For any $\xi \in X^{\mathbf{Z}^d}$, $\Gamma^c_{n,\xi}$ obeys the following large deviation principle: for every $A \subset \mathcal{M}_s(X^{\mathbf{Z}^d})$*

$$-\inf_{\mu \in \mathring{A}} I^c(\mu) \leqslant \liminf_n \frac{1}{n^{d+1}} \log \Gamma^c_{n,\xi}(A) \leqslant$$

$$\leqslant \limsup_n \frac{1}{n^{d+1}} \log \Gamma^c_{n,\xi}(A) \leqslant -\inf_{\mu \in \bar{A}} I^c(\mu)$$

where the rate function $I^c(\cdot)$ is defined by

$$I^c(\mu) = \inf\{H^c(Q): \pi(Q) = \mu\} \tag{15}$$

$\pi(Q)$ being the marginal of Q at any time t.

In (15) the function $I^c(\mu)$ is defined as an infimum, but it is actually a minimum, as shown in the following proposition.

PROPOSITION 4.1. *Let $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$ be such that $I^c(\mu) < \infty$. Then there exists $Q \in \mathcal{M}_s(\Omega)$ such that $\pi(Q) = \mu$ and $H^c(Q) = I^c(\mu)$.*

PROOF. By definition, there exists a sequence $Q_n \in \mathcal{M}_s(\Omega)$ such that $\pi(Q_n) = \mu$ and $H^c(Q_n) \to I^c(\mu)$. Since H^c has compact level sets the sequence Q_n has a limit point Q, and clearly $\pi(Q) = \mu$. By the lower semicontinuity of H^c the equality $H^c(Q) = I^c(\mu)$ easily follows. ■

We conclude this section by giving the main property of $I^c(\mu)$.

THEOREM 5. *For any $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$, $I^c(\mu) \geqslant 0$ and $I^c(\mu) = 0$ if and only if it is an invariant measure for the semigroup generated by L^c.*

PROOF. It is an obvious consequence of Proposition 4.1 and Theorem 3. ■

5. The case of reversible systems.

In this section we attack the problem of minimizing H^c under the constraint $\pi(Q) = \mu$, where $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$. In order for our argument to work we impose a quite restrictive assumption on the generator L^c. Moreover we will be able to find $I^c(\mu)$ only for μ belonging to a dense subset of $\mathcal{M}_s(X^{\mathbf{Z}^d})$. The formula we get suggests a natural conjecture on what $I^c(\mu)$ should be for a general $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$, but we have not been able to prove it.

For $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$ we denote by $\mu(\sigma(i)\,|\,\sigma)$ the conditional probability of $\sigma(i)$ with respect to the σ-field generated by the spins $\sigma(j), j \neq i$. These conditional probabilities are often called the *local specifications* of the measure μ. It is well known that the local specifications do not necessarily determine μ uniquely; when uniqueness fails we say that a *phase transition* occurs.

DEFINITION 5.1. *The operator L^c is said to be reversible with respect to the probability measure $\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$ if, for every $i \in \mathbf{Z}$, $\sigma \in X^{\mathbf{Z}^d}$*

$$c(\theta_i \sigma)\mu(\sigma(i)\,|\,\sigma) = c(\theta_i \sigma^i)\mu(-\sigma(i)\,|\,\sigma). \tag{16}$$

By translation invariance, equality (16) for $i = 0$ implies all the others. Since $c(\cdot)$ is strictly positive and local it follows easily from (16) that $\mu(\sigma(i)\,|\,\sigma)$ is also strictly positive and local. We therefore introduce the set $\mathcal{G} \subset \mathcal{M}_s(X^{\mathbf{Z}^d})$ defined by

$$\mathcal{G} = \{\mu \in \mathcal{M}_s(X^{\mathbf{Z}^d}) \colon \mu(\sigma(0)\,|\,\sigma) > 0 \text{ and local } \}.$$

It is proved by using standard argument ([6]) that, if $\mu \in \mathcal{G}$ then

$$\mu(\sigma(0)\,|\,\sigma) = \frac{1}{Z(\sigma)} \exp\left[\sum_{0 \in V \in \mathbf{Z}} J_V \sigma_V\right]$$

where

$$\sigma_V = \prod_{i \in V} \sigma(i),$$

$\{J_V\}_{V \subset \mathbf{Z}} \subset \mathbf{R}$ are such that $J_{V+i} = J_V$ for every $V \subset \mathbf{Z}$, $i \in \mathbf{Z}$, $J_V = 0$ if the diameter of V is large enough, and $Z(\sigma)$ is a normalization factor independent of $\sigma(0)$. The set $\{J_V\}_{V \in \mathbf{Z}}$ is called an *interaction*, and μ is said to be a *Gibbs measure* for that interaction. Notice that the fact that $J_V = 0$ for $\operatorname{diam}(V)$ large enough is equivalent to the locality of $\mu(\sigma(0)\,|\,\sigma)$. This is usually expressed by saying that the interaction has *finite range*.

It follows from the argument above that if L^c is reversible with respect to μ, then μ is Gibbsian for a finite range interaction. Sometimes we say that μ is a reversible measure for the system, or that the system is reversible with respect to μ. The fundamental property of reversible measures is given in the following proposition, whose proof can be found in [6].

PROPOSITION 5.2. *If L^c is reversible with respect to μ, then μ is invariant for the semigroup e^{-tL^c}. Moreover $\nu \in \mathcal{M}_s(X^{Z^d})$ is also invariant if and only if, for every $\sigma \in X^{Z^d}$,*

$$\nu((\sigma(0) \mid \sigma) = \mu(\sigma(o) \mid \sigma).$$

In the remaining part of this section we want to solve the following problem: given $\nu \in \mathcal{G}$, find the minimum of $H^c(Q)$ under the constraint $\pi(Q) = \nu$. In other words we want to compute $I^c(\mu)$ for $\mu \in \mathcal{G}$.

In what follows we assume that L^c is reversible for $\mu \in \mathcal{G}$. For a given $\nu \in \mathcal{G}$ we define

$$c^{\nu}(\sigma) = c(\sigma)\left[\frac{\mu(\sigma(0) \mid \sigma)}{\mu(-\sigma(0) \mid \sigma)} \, \frac{\nu(-\sigma(0) \mid \sigma)}{\nu(\sigma(0) \mid \sigma)}\right]^{1/2}. \tag{17}$$

Clearly c^{ν} is strictly positive and local.

PROPOSITION 5.3. *The operator $L^{c^{\nu}}$ is reversible with respect to ν.*

PROOF. It is just an elementary computation:

$$c^{\nu}(\sigma)\,\nu(\sigma(0) \mid \sigma) - c^{\nu}(\sigma^0)\,\nu(-\sigma(0) \mid \sigma) =$$

$$= \left[\frac{\nu(\sigma(0) \mid \sigma)\nu(-\sigma(0) \mid \sigma)}{\mu(\sigma(0) \mid \sigma)\mu(-\sigma(0) \mid \sigma)}\right]^{1/2} [c(\sigma)\,\mu(\sigma(0) \mid \sigma) - c(\sigma^0)\,\mu(-\sigma(0) \mid \sigma)] = 0. \quad \blacksquare$$

It follows from Propositions 5.2 and 5.3 that there is a Markovian measure Q^{ν} which is an element of $\mathcal{M}_s(\Omega)$, its Markov family of conditional distributions is $\{P_{0,\xi}^{c^{\nu}}\}$ and $\pi(Q^{\nu}) = \nu$.

Now let us fix $\omega' \in \Omega$ and $n \in N$. In the following formula any occurrence of $\omega(i)$, $i \notin \{0, 1,\dots, n-1\}$ is replaced by $\omega'(i)$. Define

$$Z_{n,\omega'}(\omega) = \exp\left\{\sum_{i=0}^{n-1}\left[\int_0^n [c(\theta_i\,\omega_t) - c^{\nu}(\theta_i\,\omega_t)]\,dt + \int_0^n \log\frac{c^{\nu}(\theta_i\,\omega_{t^-})}{c(\theta_i\,\omega_{t^-})}\,dN_t(i)\right]\right\}.$$

Now we state two lemmas that will be used later. Their proof is not given since it is a straightforward adaptation of the proofs of Lemma 4.3 in [3] and Lemma 7.1 in [2].

LEMMA 5.4. *For every* $\omega \in \Omega$, $n \in N$ *there are* $\omega_n'(\omega), \omega_n''(\omega) \in \Omega$ *such that*

$$Z_{n,\,\omega_n'(\omega)}(\omega) \leqslant \frac{dP^{c^\nu}_{0,\,\omega_0}}{dP^{c}_{0,\,\omega_0}}\Bigg|_{\mathcal{F}_n}(\omega) \leqslant Z_{n,\,\omega_n''(\omega)}(\omega).$$

LEMMA 5.5. *Let* Σ *be the set of all maps from* Ω *to* Ω. *Then, for every* $Q \in \mathcal{M}_s(\Omega)$ *such that* $H^c(Q) < \infty$

$$\lim_{n\to\infty} \frac{1}{n^2} E^Q\{\sup_{\omega''\in\Sigma} \log Z_{n,\,\omega''}(\omega) - \inf_{\omega'\in\Sigma} \log Z_{n,\,\omega'}(\omega)\} = 0\,.$$

Now we are ready to prove our main result.

THEOREM 6. *The measure* Q^ν *has the following property:*

$$H^c(Q^\nu) = \inf\{H^c(Q)\colon \pi(Q) = \nu\} = I^c(\nu)\,.$$

Moreover

$$(18)\quad I^c(\nu) = E^\nu\{c(\sigma) - c^\nu(\sigma)\} = E^\nu\left\{c(\sigma)\left[1 - \sqrt{\frac{\mu(\sigma(0)\,|\,\sigma)\nu(-\sigma(0)\,|\,\sigma)}{\mu(-\sigma(0)\,|\,\sigma)\nu(\sigma(0)\,|\,\sigma)}}\right]\right\}.$$

PROOF. By Theorem 3, Lemmas 5.4 and 5.5 we have

$$(19)\qquad H^c(Q^\nu) = \lim_{n\to\infty} \frac{1}{n^2} E^{Q^\nu}\{\log Z_{n,\,\omega'}(\omega)\}$$

for every $\omega' \in \Omega$. In particular we can choose ω' to be time-independent. Now define

$$h^\mu(\sigma) = \log \mu(\sigma(0)\,|\,\sigma) = \sum_{V\ni 0} J_V \sigma_V + \phi(\sigma)$$

where $\phi(\sigma)$ does not depend on $\sigma(0)$. Moreover

$$h^\mu(\theta_i\sigma) = \log \mu(\sigma(i)\,|\,\sigma) = \sum_{V\ni i} J_V \sigma_V + \phi(\theta_i \sigma)\,.$$

We also notice that

$$\log \frac{\mu(-\sigma(i)\,|\,\sigma)}{\mu(\sigma(i)\,|\,\sigma)} = -2\sum_{V \ni i} J_V \sigma_V .$$

Now, if we define

$$h_n^{\mu}(\sigma) = \sum_{V \cap \{0,\dots,n-1\} \neq \emptyset} J_V \sigma_V$$

then, for $i = 0,\dots, n-1$

$$h_n^{\mu}(\sigma^i) - h_n^{\mu}(\sigma) = \log \frac{\mu(-\sigma(i)\,|\,\sigma)}{\mu(\sigma(i)\,|\,\sigma)}$$

and, moreover,

$$\|h_n^{\mu}\|_{\infty} \leqslant Cn$$

for some constant $C > 0$. In an analogous way we define $h^{\nu}(\sigma)$ and $h_n^{\nu}(\sigma)$. For ω' time-independent we have

$$\sum_{i=0}^{n-1} \int_0^n \log \frac{c^{\nu}(\theta_i \omega_{t^-})}{c(\theta_i \omega_{t^-})} dN_t(i) =$$

$$= \frac{1}{2}\sum_{i=0}^{n-1}\left\{ \int_0^n [h_n^{\nu}(\omega_{t^-}^i) - h_n^{\nu}(\omega_{t^-})] dN_t(i) - \int_0^n [h_n^{\mu}(\omega_{t^-}^i) - h_n^{\mu}(\omega_{t^-})] dN_t(i) \right\} =$$

$$= \frac{1}{2}[h_n^{\nu}(\omega_n) - h_n^{\nu}(\omega_0) - h_n^{\mu}(\omega_n) + h_n^{\mu}(\omega_0)] = O(n).$$

Therefore it follows from (19) that

$$H^c(Q^{\nu}) = E^{\nu}[c(\sigma) - c^{\nu}(\sigma)].$$

We notice that the argument we have been following proves a little more than that, namely that for $Q \in \mathcal{M}_s(\Omega)$ with $H^c(Q) < \infty$

$$\lim_{n\to\infty} \frac{1}{n^2} E^Q \left\{ \log \frac{dP_{0,\omega_0}^{c^{\nu}}}{dP_{0,\omega_0}^{c}} \,|\, \mathcal{F}_n \right\} = E^Q[c(\sigma) - c^{\nu}(\sigma)].$$

Now we let Q be such that $\pi(Q) = \nu$ and $H^c(Q) < \infty$. In particular

$Q^p_\omega|_{\mathcal{F}_n} \ll P^c_{0,\omega_0}|_{\mathcal{F}_n}$. Moreover $P^{c^\nu}_{0,\omega_0}|_{\mathcal{F}_n} \approx P^c_{0,\omega_0}|_{\mathcal{F}_n}$. Therefore

$$H^c(Q) = \lim_{n\to\infty} \frac{1}{n^2} E^Q\left\{\log \frac{dQ^p_\omega}{dP^c_{0,\omega_0}} \,|\mathcal{F}_n\right\} \geq$$

$$\geq \liminf_{n\to\infty} E^Q\left\{\log \frac{dQ^p_\omega}{dP^{c^\nu}_{0,\omega_0}} \,|\mathcal{F}_n\right\} + \lim_{n\to\infty} \frac{1}{n^2} E^Q\left\{\log \frac{dP^{c^\nu}_{0,\omega_0}}{dP^c_{0,\omega_0}} \,|\mathcal{F}_n\right\} \geq$$

$$\geq \lim_{n\to\infty} \frac{1}{n^2} E^Q\left\{\log \frac{dP^{c^\nu}_{0,\omega_0}}{dP^c_{0,\omega_0}} \,|\mathcal{F}_n\right\} = E^\nu\{c(\sigma) - c^\nu(\sigma)\} = H^c(Q^\nu)$$

which completes the proof. ■

It is immediately noticed that formula (18) makes sense for every $\nu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$, and it is natural to conjecture that it actually *holds* for every $\nu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$. However we do not know how to prove it since, although $\mathcal{G}$ is dense in $\mathcal{M}_s(X^{\mathbf{Z}^d})$, a lower semicontinuous function is not uniquely determined by its restriction to a dense set. Nevertheless, density and semicontinuity, together with Theorem 6 immediately give the following.

COROLLARY 5.6. *For every* $\nu \in \mathcal{M}_s(X^{\mathbf{Z}^d})$

$$0 \leq I^c(\nu) \leq E^\nu(c(\sigma)) \leq \|c\|_\infty .$$

REFERENCES

[1] P. DAI PRA, *A stochastic control approach to reciprocal diffusion processes*, Appl. Math. Optimiz. **23** (1991), pp. 313-329.

[2] P. DAI PRA, *Space-time large deviations for interacting particle systems*, Comm. Pure Appl. Math., **46** (1993), pp. 387-422.

[3] P. DAI PRA, *Large deviations and equilibrium measures for stochastic spin systems*, Stochastic Processes and Their Applications, 48 (1993), pp. 9-30.

[4] S. N. ETHIER and T. G. KURTZ, *Markov processes, characterization and convergence*, John Wiley & Sons (1986).

[5] H. FOLLMER, *Random fields and diffusion processes*, in *Ecole d'Eté de Saint Flour*, XV-XVII (1985-87), p. 101-203, Lecture Notes in Mathematics, **1362**, Springer-Verlag (1988).

[6] T. M. LIGGETT, *Interacting Particle Systems*, Springer-Verlag (1988).

[7] M. NAGASAWA, *Transformation of diffusion and Schrödinger processes*, Prob. Th. Rel. Fields, **82** (1989), pp. 109-136.

[8] M. PAVON - A. WAKOLBINGER, *On free energy, stochastic control and Schrödinger processes*, in *Proc. Workshop on Modeling and Control of Uncertain Systems*, edited by G. B. DI MASI, A. GOMBANI and A. KURZHANSKI, Birkhauser (1991).

[9] S. R. VARADHAN, *Large Deviations and Applications, CBMS-NSF Regional Conference series in Applied Mathematics*, Vol. 46, Society for Industrial and Applied Mathematics, Philadelphia (1984).

[10] A. WAKOLBINGER, *A simplified variational characterization of Schrödinger processes*, J. Math. Phys, **30** (12) (1989), pp. 2943-2946.

[11] J. C. ZAMBRINI, *Stochastic mechanics according to Schrödinger*, Phys. Rev. A, **33** (3) (1986), pp. 1532-1548.

Manoscritto pervenuto in redazione il 26 novembre 1992
e, in forma definitiva, il 18 febbraio 1993.