Monotonicity of the Free Energy in the Stochastic Heisenberg Model*

William D. Wick**

Mathematics Department, University of Washington, Seattle, WA 98195, USA

Abstract. The specific free energy of the state at time t of the stochastic Heisenberg model is shown to be non-increasing with t, and to strictly decrease whenever this state is not a Gibbs state of the Hamiltonian. The initial state is assumed to be translation invariant and suitably smooth. For such states a convergence theorem is obtained.

I. Introduction

The classical Heisenberg model is one of a class of lattice spin models in which the range of a single spin is a sphere S^n, $n \geq 1$, rather than the two-point set S^0, as in the Ising model. The Hamiltonian for these models is given (formally) by:

$$H = -\tfrac{1}{2} \sum_{\substack{x, y \in L \\ |x-y|=1}} \xi(x) \cdot \xi(y), \tag{1}$$

where $\xi(x) \in S^n$ is the spin at site x, L is a d-dimensional lattice, and the "$\cdot$" denotes dot product in R^{n+1}. Special cases include the planerotor models ($n=1$) and the classical Heisenberg model ($n=2$, and usually $d=3$).

For these models with continuous symmetry groups invariance of phase is expected if $d=2$ and phase-transition with associated continuous symmetry breaking if $d \geq 3$. These facts were established in [1, 2].

The stochastic Heisenberg model, a probabilistic model for the dynamics of the classical Heisenberg model, was introduced by Faris in [3]. This is a Markovian model with infinitesimal generator

$$\Omega = \Delta - \beta \nabla H \cdot d. \tag{2}$$

In this equation the first term is the infinite-dimensional Laplace operator and generates a Brownian motion of the individual spins. The second term is the inverse

---

temperature (β) times the gradient of $(-H)$, forming a vector field X which generates a deterministic drift. The Markovian diffusion process generated by Ω has the Gibbs states associated with H as invariant, or equilibrium, measures. More precisely the Gibbs states of H form the set of reversible states for this process.

The existence of this Markov process was established in [3], and a high-temperature uniqueness theorem for the equilibrium state was proved in [4]. In [5] the process was shown to converge, at an exponential rate, to the (unique) equilibrium state at high temperatures. ([5] also contains a discussion of reversible states and of the global Markov property.) These developments parallel earlier work on the stochastic Ising model.

It is not presently known whether uniqueness of the Gibbs state (absence of phase transition) implies convergence to equilibrium of the Markov process. (To the author's knowledge it is not even known whether non-Gibbs invariant states exist for either the stochastic Ising or Heisenberg models.) However if we impose additional symmetry and regularity conditions on the initial state, we can prove convergence to equilibrium whenever the Gibbs state is unique. If non-uniqueness holds, convergence occurs to the *set* of Gibbs states. We will exploit the properties of a functional on the set of states called the specific free energy (or free energy per spin). This functional plays the role of a Liapunov function: it is monotonically decreasing in time and reaches a minimum on the Gibbs states.

We will work in the same generality as [5]. We take the spin space M to be a compact, orientable C^∞ Riemannian manifold with metric tensor g and volume measure σ (normalized to be a probability measure). It will be convenient to assume that $L=(Z+\frac{1}{2})^d$, the lattice of points in R^d obtained by translating the integer lattice Z^d by the vector $(\frac{1}{2}, \ldots, \frac{1}{2})$. The state space of our Markov process is $\Xi = M^L$. The Hamiltonian will be given (formally) by

$$H = -\sum_{F \subset L} J(F), \tag{3}$$

where $\{J(F) : F \subset L, F \text{ finite}\}$ forms a family of potential functions satisfying:

(i) Smoothness: $J(F)$ is a C^∞ function of $\xi(x)$, $x \in F$.

(ii) Translation invariance: $\tau_x J(F) = J(\tau_x F)$ for all $x \in Z^d$, where $\tau_x f(\xi) = f(\tau_x^{-1} \xi)$ for a function f on Ξ, and $\tau_x \xi(y) = \xi(y-x)$ is the induced action on Ξ of translation by x.

(iii) Finite-range condition: $J(F) = 0$ if diam $F > r$, where r (the range of the interaction) is a fixed positive number, and the distance on L is given by: dist $(x, y) = \sum_i |x_i - y_i|$.

Although (3) does not define H as a function on Ξ, the expression

$$\nabla H \cdot d = \sum_{x \in L} \nabla_x H \cdot d_x = \sum_{x \in L} d_x H \cdot g^{-1} \cdot d_x$$

[where d_x is the differential with respect to $\xi(x)$] yields a well-defined differential operator with domain the C^∞ functions of finitely many manifolds. Equation (2) then defines Ω on this domain.

From (ii) and (iii) it is easily seen that there exist finite constants K_1 and K_2 with

$$\|V_x H\| \leq K_1, \qquad \|V_x V_y H\| \leq K_2,$$

for all x and y in L, where we use V_x also for the covariant derivative with respect to $\xi(x)$, and the norms are supremums over $\varXi$ of the magnitudes of the tensors. Assuming these bounds and the finite-range condition, Faris in [3] proved that the closure of $\varOmega$ generates a Markov semigroup on $C(\varXi)$. (We denote the closure of $\varOmega$ again by $\varOmega-$ a convention we adopt throughout the paper for closable operators. Dissipative operators are automatically closable.) Condition (ii) implies additionally that $S(t)$ commutes with τ_x for all $t \geq 0$ and $x \in Z^d$.

We now define the free energy functional. Let v be a state (probability measure) on $\varXi$. We define the energy inside a finite set $\varLambda \subset L$ by:

$$H_\varLambda = - \sum_{F \subset \varLambda} J(F),$$

and the average energy inside $\varLambda$ of v by: $v(H_\varLambda) = \int H_\varLambda dv$. We define the entropy inside $\varLambda$ of v (with respect to the product measure σ) by:

$$S_\varLambda(v) = \begin{cases} -\int \varPhi(d(v|\varLambda)/d\sigma)d\sigma, & \text{if } v|\varLambda \ll \sigma|\varLambda; \\ -\infty, & \text{otherwise}. \end{cases}$$

Here $v|\varLambda$ denotes v restricted to the σ-algebra generated by $\xi(x)$ for $x \in \varLambda$, and $\varPhi(t) = t \log t - t + 1$ for $t > 0$ and $\varPhi(0) = 1$. (For future reference, note that $\varPhi$ is non-negative and continuous for $0 \leq t < \infty$ and differentiable with derivative $\log t$ for $t > 0$.)

The free energy $F_\varLambda(v)$ of v inside $\varLambda$ is defined to be:

$$F_\varLambda(v) = v(H_\varLambda) - TS_\varLambda(v)$$

$(T = \beta^{-1})$. The specific free energy of v is then defined as

$$f(v) = \limsup_{\varLambda \uparrow S} |\varLambda|^{-1} F_\varLambda(v), \tag{4}$$

where the limit is taken along a sequence of finite boxes $\varLambda$ which increase to L "in the sense of van Hove" [7]; that is, $|\varLambda| \to \infty$ in such a way that $|\varLambda|^{-1}|\varLambda_r^o| \to 1$ for any positive number r. (Here $|\varLambda|$ denotes the cardinality of $\varLambda$ and $\varLambda_r^o$ is the "r-interior" of $\varLambda$; i.e. the set $\{x \in \varLambda : \text{dist}(x, \partial \varLambda) > r\}$.) If H has finite-range potentials and both H and v are translation invariant [v is translation invariant if $v(\tau_x f) = v(f)$ for all $f \in C(\varXi)$ and $x \in Z^d$] it can be shown that the superior limit in (4) is actually a limit. In this case we have also an infinite-volume version of the Gibbs Variational Principle: If μ is translation invariant and $f(\mu) = \inf f(v)$, the infimum being over the translation invariant states, then μ is a Gibbs state of H. For discussion of these ideas see [7], especially Chap. 7.

For the stochastic Ising model Holley in [8] proved that the specific free energy of the state at time t is non-increasing with t, and if the initial state was translation invariant, strictly decreases whenever the state in not a Gibbs state of

H. It follows in the translation invariant case that the state at time t approaches the set of Gibbs states as $t \to \infty$. Thus uniqueness of the Gibbs state gives a convergence theorem. Holley's theorem was generalized to include long-range interactions by Higuchi and Shiga [9] and to more general jump processes by Sullivan [10]. Using an improved expression for the time-derivative of the free energy due to Moulin Ollagnier and Pinchon [11], Holley and Stroock in [12] proved that for $d \leq 2$ every equilibrium state of the stochastic Ising model is a Gibbs state.

   Our approach will be almost identical to Holley's. The single essential difference is that we have to restrict further the class of initial states. We shall call a state ν *locally smooth* if

   (i) $\nu|\Lambda \ll \sigma|\Lambda$ for all finite $\Lambda \subset L$;
   (ii) $q_\Lambda = d(\nu|\Lambda)/d\sigma$ is a C^1, positive function and there exists a constant C such that

$$\int |V_x q_\Lambda| d\sigma \leq C$$

for all x and Λ.

   The class of locally smooth states is fairly large. For instance, it includes all Gibbs states of Hamiltonians $\tilde{H}$ for which $V_x \tilde{H}$ is uniformly bounded. We restrict all initial states to be locally smooth. The purpose of the theorems in Sect. II is essentially to allow us to prove that this class is invariant under the time evolution.

   The following theorems will be proved.

I.1. Theorem. *Let ν be a locally smooth initial state and $\nu(t) = \nu S(t)$ the state at time t. Then the specific free energy $f(\nu(t))$ is non-increasing with t. If in addition ν is translation invariant, then $f(\nu(t))$ strictly decreases over a time interval $[t_0, t_0 + \varepsilon)$ for some $\varepsilon > 0$ whenever $\nu(t_0)$ is not a Gibbs state of the Hamiltonian H.*

Corollary. *The only locally smooth and translation invariant equilibrium states are Gibbs states.*

I.2. Theorem. *Let ν be a translation invariant, locally smooth initial state with $f(\nu) < \infty$. If μ is a weak limit point of $\{\nu(s): s \geq t\}$ for each $t \geq 0$, then μ is a Gibbs state of H.*

Corollary. *If ν is as above and the Gibbs state μ of H is unique, then $\nu(t) \to \mu$ weakly as $t \to \infty$.*

Remark. As another corollary of Theorem 12 we obtain the variational principle mentioned previously for locally smooth states.

   The proof of these theorems has been divided into several sections. In Sect. II a smoothness-preserving property of the time evolution operator is established. In Sect. III a formula for the rate of decrease of the specific free energy is derived. As a guide to the reader we list the essential steps in this derivation.
Steps (1) and (2): The infinite-volume time evolution is replaced by a limit of finite-volume evolutions. Translation-invariance is restored by imposing periodic boundary conditions on the potentials.
Steps (3) through (6): The time derivative of the free energy in a box Λ is computed for the time evolution in a box $\Gamma_n \supset \Lambda$. Boundary terms are shown to be

of order $|\partial_r \Lambda_r^o|$, where for $F \subset L$, $\partial_r F = \{y \in L \backslash \Lambda : \mathrm{dist}(y, F) \leqq r\}$. The remaining terms are non-positive. Taking the limits $\Gamma_n \uparrow L$, then $\Lambda \uparrow L$, the boundary terms drop out and we obtain a bound on the time-derivative of $f(v(t))$.

Steps (7) and (8): Imposing translation invariance on the initial state, the bound on the time-derivative of the free energy in Λ is shown to be monotonic in Λ. This yields an improved formula for the rate of decrease of $f(v(t))$.

In Sect. IV we prove the theorems with the aid of a proposition characterizing the Gibbs states of the Hamiltonian by a set of partial differential equations.

After completing this manuscript the author received paper [14], in which the authors obtain analytically better results for the planerotor models. They prove a smoothness-*improving* property of the time evolution and are thus able to drop the local smoothness assumption on the initial state as well as the condition $f(v) < \infty$.

II. Smoothness-Preserving Properties of the Time Evolution

In this section we employ methods similar to those used by Faris in [3] to prove that the semigroup associated with our diffusion process preserves certain classes of smooth functions on the configuration space. We shall work on finite-dimensional manifolds, but the bounds proved will be independent of the dimension of the manifold. The theory elaborated herein may be regarded as an addendum to Faris' existence theorem [3].

Let (N, g) be a finite-dimensional, C^∞ Riemannian manifold. We will use an index notation for the values of tensors on N with upper indices representing contravariant values and lower indices, covariant ones. Thus X^a is a vector, u_a is a one-form, etc. Indices are raised or lowered with the metric tensor g and repeated indices imply contraction. The covariant derivative will be written as V_a. [Note: on scalar functions f, $V_a f = d_a f$ will denote the differential of f (a one-form); the gradient of f will be written $V^a f = g^{ab} V_b f$.] The magnitude of a tensor T at a point of N will be calculated with the metric and written: $|T|$. Thus e.g. $|X| = (X^a X_a)^{1/2}$ for a vectorfield X.

Let H be a smooth function on N, $X^a = -\beta V^a H$, and $\Omega = \Delta + X \cdot d$ with domain $C^\infty(N)$. The semigroup $\{S(t) : t \geqq 0\}$ generated by the closure of Ω has a unique reversible measure $\mu = Z^{-1} \exp(-\beta H)$ cf. [5, Sect. V]. We shall be interested in the Banach space $L^1(N, T^1, \mu)$ of measurable one-forms on N with norm

$$\|u\|_1 = \int_N |u| d\mu.$$

We shall need in addition the operator $\tilde{\Omega}$ with domain $C^\infty(N, T^1)$, the space of smooth one-forms on N, defined by:

$$(\tilde{\Omega} u)_a = V^b V_b u_a + X^b V_b u_a.$$

Faris proved in [3] that the closure of $\tilde{\Omega}$ generates a c_0-contraction semigroup on $C(N, T^1)$, the space of continuous one-forms on N. (c_0 means strongly continuous in t.) This process had been discussed originally by Ito [6].

II.1. Theorem. $\tilde{\Omega}$ *generates a c_0-contraction semigroup on $L^1(N, T^1, \mu)$.*

Remark. Defining $L^p(N, T^1; \mu)$ similarly with norm

$$\|u\|_p = \left(\int_N |u|^p d\mu\right)^{1/p},$$

one can extend this theorem to L^p for $1 \leq p < \infty$.

II.2. Theorem. *Let $f \in C^\infty(N)$. There is a bound, for all $t \geq 0$:*

$$\|dS(t)f\|_1 \leq e^{\bar{C}t}\|df\|_1$$

with $\bar{C}$ a positive constant equal to $\|\bar{R}\| + \beta\|\nabla\nabla H\|$, where $\bar{R}$ is the Ricci tensor on N, and the norms are the supremums over N of the norms of the linear transformations $\bar{R}_a^b$ and $\nabla_a\nabla^b H$. (Alternatively one could take the magnitudes of the tensors.)

When N is a product of Riemannian manifolds, $N = M^A$ with A a finite set, it will be convenient for the developments in the next chapter to consider a different Banach space of one-forms. For this purpose and for clarity of exposition we introduce projections π_x, $x \in A$ which project tensor indices in the direction of the tangent or contangent space to the manifold M_x. Thus if X is a vector field on N, $\pi_x X$ is the vector field on N obtained from X by projecting $X(\xi)$ into the tangent space to M_x at $\xi(x)$ (and similarly for one-forms and tensor fields). We then define $L_A^1(M^A, T^1, \mu)$ to be the Banach space of measurable one-forms with norm

$$\|u\|_{1, A} = \max_{x \in A} \int_N |\pi_x u| d\mu.$$

II.3. Theorem. $\tilde{\Omega}$ *generates a c_0-contraction semigroup on $L_A^1(M^A, T^1, \mu)$.*

II.4. Theorem. *If f is in $C^\infty(M^A)$ there is a bound, for all $t \geq 0$:*

$$\|dS(t)f\|_{1, A} \leq e^{Ct}\|df\|_{1, A},$$

where $C = \|R\| + \max_x \sum_y \beta\|\pi_x \nabla \pi_y \nabla H\|$, R is the Ricci tensor on M, and the norms are the supremums of the magnitudes of the tensors.

Proof of Theorem II.1. Clearly $\tilde{\Omega}$ is densely defined and, by Faris' theorem and the density of $C(N, T^1)$ in L^1, $I - \kappa\tilde{\Omega}$ has dense range for all $\kappa \geq 0$. Thus we need only show that $\tilde{\Omega}$ is dissipative. For the proof of this we need the following criterion: If A is an operator on a Banach space B, then A is dissipative if for every u in $D(A)$ there exists a normalized tangent functional u^* to u (that is, an element $u^* \in B^*$, the dual of B, of norm one with $\langle u^*, u \rangle = \|u\|$) such that $\langle u^*, Au \rangle \leq 0$. In fact then $\|(I - \kappa A)u\| \geq \langle u^*, (I - \kappa A)u \rangle \geq \|u\|$.

Let $u \in C^\infty(N, T^1)$ and define a measurable vector field on N by setting $Y^a = |u|^{-1}u^a$ on the set E where u doesn't vanish, and set $Y = 0$ outside E. Define a linear functional on L^1 by: $\langle Y, u \rangle = \int Y^a u_a d\mu$. It is easily checked that this defines a normalized tangent functional to u and we will show that $\langle Y, \tilde{\Omega}u \rangle \leq 0$.

We have

$$\langle Y, \tilde{\Omega}u \rangle = \int Y^a \tilde{\Omega}u_a d\mu = \int |u|^{-1}u^a\tilde{\Omega}u_a d\mu$$

(where the integral is restricted to E). Since $\mu = \varrho\sigma$ with ϱ smooth we have, by an application of the Leibniz rule,

$$
\begin{aligned}
\int |u|^{-1} u^a \tilde{\Omega} u_a d\mu &= \int d\sigma \varrho |u|^{-1} u^a \{V^b V_b + X^b V_b\} u_a \\
&= \int d\sigma \{V^b(\varrho|u|^{-1} u^a V_b u_a) + |u|^{-1} u^a V_b u_a(-V^b \varrho + \varrho X^b) \\
&\quad - \varrho V^b(|u|^{-1} u^a) \cdot V_b u_a\} .
\end{aligned}
\tag{1}
$$

The second term on the right side of (1) vanishes since $-V\varrho + \varrho X \equiv 0$. For the third term note that on the set E where u doesn't vanish,

$$
\begin{aligned}
V^b(|u|^{-1} u^a) \cdot V_b u_a &= V^b(|u|^{-1} u^a) \cdot V_b(|u|^{-1} u_a \cdot |u|) \\
&= |V(|u|^{-1} u)|^2 \cdot |u| + \tfrac{1}{2} V^b(|u|^{-1} u_a |u|^{-1} u^a) \\
&= |V(|u|^{-1} u)|^2 \cdot |u| + 0 .
\end{aligned}
$$

Thus the third term makes a negative contribution to the integrand.

This leaves the first term, to which we apply the divergence theorem, as follows: First shrink E to $E_\varepsilon = \{|u| \geq \varepsilon\}$ for $\varepsilon > 0$. Note that

$$
\begin{aligned}
\int\limits_{E_\varepsilon} d\sigma V^b(\varrho|u|^{-1} u^a V_b u_a) &= \int\limits_{E_\varepsilon} d\sigma V^b(\tfrac{1}{2}\varrho|u|^{-1} V_b |u|^2) \\
&= \int\limits_{E_\varepsilon} d\sigma V^b(\varrho V_b |u|) ,
\end{aligned}
$$

so that by the divergence theorem the latter equals

$$
\int\limits_{\partial E_\varepsilon} ds \varrho V_b |u| \cdot n^b ,
$$

where n is the unit outward normal vector to ∂E_ε and ds is the surface measure on ∂E_ε. Since $V_b|u|$ points inward, the integrand is non-positive. Letting $\varepsilon \to 0$ we obtain the first term in (1) which thus also makes a non-positive contribution to the integral. Therefore $\tilde{\Omega}$ is dissipative. $\square$

Proof of Theorem II.2. We begin with the equation intertwining Ω, $\tilde{\Omega}$, and the differential operator d used by Faris in proving his existence theorem:

$$
d_a \Omega f = \tilde{\Omega} d_a f + (\bar{R}_a^b + V_a X^b) d_b f ,
$$

valid for any $f \in C^\infty(N)$. Here $\bar{R}$ is the Ricci tensor of (N, g). We rewrite this equations as

$$
d\Omega f = (\tilde{\Omega} + \tilde{P}) df .
\tag{2}
$$

The closure of $\tilde{\Omega}$ generates a c_0-contraction semigroup by Theorem II.1. $\tilde{P}$ represents a bounded perturbation:

$$
\begin{aligned}
\|\tilde{P}u\|_1 &= \int |(\bar{R} + VX)u| d\mu \\
&\leq (\|\bar{R}\| + \|VX\|) \int |u| d\mu ,
\end{aligned}
$$

where the norms on the tensors represent supremums over N of the operator norms. (Alternatively one may take the magnitudes of the tensors calculated with the metric g-these represent the Hilbert-Schmidt norms of the linear transformations and so are larger than the operator norms.)

We shall need the following well-known proposition from semigroup theory.

Proposition 1. *Let (the closure of) $\tilde{\Omega}$ be the generator of a c_0-contraction semigroup on the Banach space $\tilde{B}$, and let $\tilde{P}$ be a bounded perturbation. Then the closure of $\tilde{\Omega}+\tilde{P}$ generates a c_0-semigroup $\{\tilde{S}(t):t\geq 0\}$ and there is a bound, for all $u\in\tilde{B}$:*

$$\|\tilde{S}(t)u\| \leq e^{\|\tilde{P}\|t}\|u\|.$$

[*Proof.* It is easily checked that the closure of $\tilde{\Omega}+\tilde{P}-\|\tilde{P}\|I$ generates a c_0-contraction semigroup $\{\bar{S}(t)\}$ and that $\bar{S}(t)$ is related to $\tilde{S}(t)$ by: $\tilde{S}(t)=e^{\|\tilde{P}\|t}\bar{S}(t)$.]

Since $(I-\kappa\Omega)^{-1}:C^\infty(N)\to C^\infty(N)$ (by elliptic regularity) we obtain from (2):

$$d(I-\kappa\Omega)^{-1}f=(I-\kappa\tilde{\Gamma})^{-1}df \tag{3}$$

for all $\kappa\geq 0$ and $f\in C^\infty(N)$, where $\tilde{\Gamma}$ is the closure of $\tilde{\Omega}+\tilde{P}$. Iterating (3) n times with $\kappa=t/n$ and letting $n\to\infty$, we obtain from the Hille-Yosida theorem

$$dS(t)f=\tilde{S}(t)df,$$

where $\{\tilde{S}(t)\}$ is generated by $\tilde{\Gamma}$. Now the proposition applies. $\square$

Proof of Theorem II.3. Again we need only prove that $\bar{\Omega}$ is dissipative on L^1_Λ. Let $u\in C^\infty(M, T^1)$ and define a measurable vector field as follows. Let $x_0\in\Lambda$ be the element for which

$$\int |\pi_{x_0}u|d\mu = \max_{x\in\Lambda}\int |\pi_x u|d\mu$$

and define $\pi_{x_0}Y^a=|\pi_{x_0}u|^{-1}\pi_{x_0}u^a$ or zero and $\pi_x Y^a\equiv 0$ for $x\neq x_0$. Again it is easily checked that Y defines a normalized tangent functional in $(L^1_\Lambda)^*$. Since π_{x_0} commutes with $\tilde{\Omega}$ we have

$$\langle Y, \tilde{\Omega}u\rangle = \int |\pi_{x_0}u|^{-1}\pi_{x_0}u^a\cdot\pi_{x_0}\tilde{\Omega}u_a d\mu$$
$$= \int |\pi_{x_0}u|^{-1}\pi_{x_0}u^a\cdot\tilde{\Omega}\pi_{x_0}u_a d\mu.$$

Thus the proof of Theorem II.1 applies with u_a replaced by $\pi_{x_0}u_a$ and we obtain $\langle Y, \tilde{\Omega}u\rangle\leq 0$. Thus $\tilde{\Omega}$ is dissipative. $\square$

Proof of Theorem II.4. The proof of Theorem II.2 applies but with a different bound on the perturbation term. We have

$$\|\tilde{P}u\|_{1,\Lambda}=\max_x\int |\pi_x\tilde{P}u|d\mu$$

$$\leq \max_x\sum_y(\|\pi_x\bar{R}\pi_y\|+\beta\|\pi_x\nabla\pi_y\nabla H\|)\cdot\int |\pi_y u|d\mu.$$

Since $\pi_x\bar{R}\pi_y=0$ if $x\neq y$ we obtain the constant used in the theorem on the right-hand side. $\square$

III. Rate of Decrease of the Free Energy

Step 1. Let $\Lambda \subset L$ be a finite set. The free energy $F_\Lambda(v)$ inside Λ is a lower semicontinuous function of v in the weak topology:

$$\liminf_{n \to \infty} F_\Lambda(v_n) \geqq F_\Lambda(v) \cdot \tag{1}$$

if $v_n \to v$ weakly. The proof of this assertion follows from the fact that H_Λ is assumed to be continuous, and from the following lemma (stated in somewhat greater generality in [10]).

Lemma. *Let $\{g_1, \ldots, g_m\}$ be a finite partition of unity on M^Λ consisting of Borel functions, and define a state v^* by:*

$$dv^*/d\sigma = \sum_{i=1}^n (\textstyle\int g_i dv)(\textstyle\int g_i d\sigma)^{-1} g_i.$$

Then $S_\Lambda(v^*) \geqq S_\Lambda(v)$. Furthermore if α is any real number with $\alpha > S_\Lambda(v)$, there exists a partition of unity $\{g_1, \ldots, g_m\}$ consisting of continuous functions on M^Λ for which $\alpha > S_\Lambda(v^*)$.

Step 2. In this step we introduce finite-volume evolution operators approximating $S(t)$. Let $\Gamma_n = [-n, n]^d \cap L$ for $n > 3r$. (r is the range of the potentials.) In order to preserve translation invariance we impose periodic boundary conditions on the potentials, as follows. Let $\{J_n(F) : F \subset \Gamma_n\}$ be the uniquely defined set of potentials which are invariant under translations $(\mathrm{mod} - 2n)$ of Γ_n and satisfy: $J_n(F) = J(F)$ if $F \subset (\Gamma_n)_r^0$. Define $H_n = - \sum_{F \subset \Gamma_n} J_n(F)$. By pulling back along the canonical projection:

$\Xi \to M^{\Gamma_n}$ we may regard functions on M^{Γ_n} as being defined on Ξ and define $\Omega_n = \Delta - \beta \nabla H_n \cdot d$. Ω_n generates a Markov semigroup $\{S_n(t)\}$ on $C(\Xi)$ leaving the space $C(M_n^\Gamma)$ invariant and commuting with translations $(\mathrm{mod}\text{-}2n)$ of this space.

The space D of C^∞ functions of finitely many spins is a common core of all Ω_n and Ω, and $\Omega_n \to \Omega$ on D. [Note that $\nabla_x H_n = \nabla_x H$ if $x \in (\Gamma_n)_r^0$.] By the Trotter-Kurtz theorem [13], $S_n(t) \to S(t)$ as $n \to \infty$ strongly on $C(\Xi)$. Hence $v S_n(t) = v_n(t) \to v(t)$ weakly. We then have from Step 1, for any finite $\Lambda \subset L$:

$$F_\Lambda(v(t)) - F_\Lambda(v) \leqq \liminf_{n \to \infty} (F_\Lambda(v_n(t)) - F_\Lambda(v))$$

$$\leqq \limsup_{n \to \infty} (F_\Lambda(v_n(t)) - F_\Lambda(v)). \tag{2}$$

Step 3. If n is sufficiently large [such that $\Lambda \subset (\Gamma_n)_r^0$] we have

$$F_\Lambda(v_n(t)) - F_\Lambda(v) = \int_0^t G_{\Lambda, n}(\tau) d\tau \tag{3}$$

with

$$G_{\Lambda, n}(t) = \int \Omega_n(H_\Lambda + T \cdot \log q_{\Lambda, n}(t)) dv_n(t), \tag{4}$$

where

$$q_{A,n}(t) = d(v_n(t)|A)/d\sigma. \tag{5}$$

To establish this formula we prove that $F_A(v_n(t))$ is differentiable for $t > 0$ with derivative $G_{A,n}(t)$ and continuous at $t = 0$. Let g be a C^2 function on M^A. Since $q_{A,n}(t)(\xi)$ is a C^∞ function of $(t, \xi) \in (0, \infty) \times M^A$ we have

$$\int g d/dt q_{A,n}(t) d\sigma = d/dt \int g q_{A,n}(t) d\sigma$$
$$= d/dt \int S_n(t) g dv$$
$$= \int \Omega_n g dv_n(t).$$

The derivative of the energy term in the free energy is therefore

$$d/dt \int q_{A,n}(t) H_A d\sigma = \int \Omega_n H_A dv_n(t),$$

and the derivative of the entropy term is

$$d/dt T \int \Phi(q_{A,n}(t)) d\sigma = T \int \Phi'(q_{A,n}(t)) d/dt q_{A,n}(t) d\sigma$$
$$= T \int \Omega_n \log q_{A,n}(t) dv_n(t).$$

To prove the continuity assertion, let $q_n(t) = d(v_n(t)|\Gamma_n)/d\sigma$. It is readily shown that $q_n(t)$ is the solution of the "forward equation"

$$d/dt q_n(t)(\cdot) = \Omega_n^t q_n(t)(\cdot)$$

with initial condition $q_n(0) = d(v|\Gamma_n)/d\sigma$. Since Ω_n^t is the formal adjoint of Ω with respect to σ, it is the sum of first and second order terms (which generate a contraction semigroup on $C(M^{\Gamma_n})$) and a zeroth order term (representing a bounded perturbation). Hence its closure generates a c_0-semigroup on $C(M^{\Gamma_n})$. Thus $q_n(t) \to q_n(0)$ uniformly as $t \to 0$. Since $q_{A,n}(t)$ is obtained from $q_n(t)$ by integrating out spins with product measure, $q_{A,n}(t) \to q_{A,n}(0)$ as well. This gives the continuity of $F_A(v_n(t))$ at $t = 0$.

Step 4. Define

$$\tilde{G}_{A,n}(t) = \int \Omega_A (H_A + T \log q_{A,n}(t)) dv_n(t), \tag{6}$$

where $\Omega_A = \Delta - \beta \nabla H_A \cdot d$. In this step we show that

$$G_{A,n}(t) = \tilde{G}_{A,n}(t) + C(t)|\partial_r A_r^0| \tag{7}$$

with $C(t)$ bounded independently of n and A.

To prove this we need uniform L^1-bounds on the differential of $q_{A,n}(t)$. We appeal to Theorem II.4 of Sect. II. Let $\mu_n = \varrho_n \sigma$ be the unique reversible measure for $\{S_n(t)\}$, cf. [5]. [ϱ_n is a constant times $\exp(-\beta H_n)$.] Writing $p_n(t) = d(v_n(t)|\Gamma_n)/d\mu_n$ it is easily seen (using the reversibility of μ_n) that $p_n(t) = S_n(t) p_n(0)$. Since $q_n(t) = p_n(t) \varrho_n$

we have

$$\sup_x \int |d_x q_n(t)| d\sigma = \sup_x \int |d_x(p_n(t)\varrho_n)| d\sigma$$

$$\leq \|dp_n(t)\|_{1,\Gamma_n} + \sup_x \int |\beta d_x H_n| dv_n(t)$$

$$\leq e^{Ct} \|dp_n(0)\|_{1,\Gamma_n} + \beta K_1$$

$$= e^{Ct} \sup_x \int |d_x(q_n(0)\varrho_n^{-1})| d\mu_n + \beta K_1$$

$$\leq e^{Ct} \left(\sup_x \int |d_x q_n(0)| d\sigma + \beta K_1 \right) + \beta K_1. \tag{8}$$

[Theorem II.4 was used in the third line with $C = \|R\| + \beta(2r+1)^d K_2$.] Since v was assumed to be locally smooth this last bound is independent of n. The same bound holds with $q_{A,n}(t)$ replacing $q_n(t)$, since the former is obtained from the latter by integrating out spins.

The difference between $G_{A,n}(t)$ and $\tilde{G}_{A,n}(t)$ is a term equal to

$$- \sum_{x \in \partial_r A_r^o} \beta \int (\nabla_x H_A - \nabla_x H) \cdot d_x(H_A + T \cdot \log q_{A,n}(t)) dv_n(t).$$

This term is bounded by

$$2\beta K_1 \left(\sum_{x \in \partial_r A_r^o} \int |d_x(H_A + T \cdot \log q_{A,n}(t))| q_{A,n}(t) d\sigma \right),$$

which in turn is less than

$$2\beta K_1 \left(\sum_{x \in \partial_r A_r^o} (K_1 + T \int |d_x q_{A,n}(t)| d\sigma) \right). \tag{9}$$

Combining the bounds (8) and (9) we obtain Eq. (7).

Step 5. Applying the divergence theorem to the integral in the definition of $\tilde{G}_{A,n}(t)$ yields

$$\tilde{G}_{A,n}(t) = \int q_{A,n}(t)(\Delta - \beta \nabla H_A \cdot d)(H_A + T \cdot \log q_{A,n}(t)) d\sigma$$

$$= - \sum_{x \in A} \int (\nabla_x q_{A,n}(t) + \beta q_{A,n}(t) \nabla_x H_A) \cdot d_x(H_A + T \cdot \log q_{A,n}(t)) d\sigma$$

$$= - \sum_{x \in A} T \int q_{A,n}(t)^{-1} |\nabla_x q_{A,n}(t) + \beta q_{A,n}(t) \nabla_x H_A|^2 d\sigma. \tag{10}$$

Step 6. Define

$$\tilde{G}_A(t) = \limsup_{n \to \infty} \tilde{G}_{A,n}(t).$$

Applying Fatou's inequality [note that $\tilde{G}_{A,n}(t) \leq 0$ by Eq. (10)] and the results of the previous steps we see that

$$F_A(v(t)) - F_A(v) \leq \int_0^t \tilde{G}_A(\tau) d\tau + \tilde{C}(t) |\partial_r A_r^o|.$$

Dividing by $|\Lambda|$, letting $\Lambda \uparrow L$ in the sense of van Hove, and applying Fatou again we obtain:

$$f(v(t)) - f(v) \leqq \int_0^t \tilde{g}(\tau) d\tau, \tag{11}$$

where

$$\tilde{g}(t) = \limsup_{\Lambda \uparrow S} |\Lambda|^{-1} \tilde{G}_\Lambda(t). \tag{12}$$

Step 7. We assume from now on that v is translation invariant. Our aim now is to obtain improved versions of formulas (10)–(12) by exploiting translation invariance. Let $\Lambda_m = [-2^m, 2^m]^d \cap L$ and define for each integer m and sufficiently large n [such that $\Lambda_m \subset (\Gamma_n)_r^o$],

$$\Pi_{m,n}(t) = -\sum_{x \in (\Lambda_m)_r^o} T \int q_{\Lambda_m, n}(t)^{-1} |\nabla_x q_{\Lambda_m, n}(t) + \beta q_{\Lambda_m, n}(t) \nabla_x H|^2 d\sigma \tag{13}$$

and

$$\Pi_m(t) = \limsup_{n \to \infty} \Pi_{m,n}(t). \tag{14}$$

From Step 5 we see that $\Pi_m(t)$ represents the rate at which the free energy in Λ_m is decreasing due to the evolution of the spins in $(\Lambda_m)_r^o$. In this step we establish the inequality:

$$\Pi_m(t) \leqq 2^d \Pi_{m-1}(t). \tag{15}$$

Since $|\Lambda_m| = 2^{(m+1)d}$ it follows that the infinite-volume limit

$$\Pi(t) = \lim_{m \to \infty} |\Lambda_m|^{-1} \Pi_m(t) = \inf_m |\Lambda_m|^{-1} \Pi_m(t)$$

exists.

To prove inequality (15) let Λ_{m-1}^i, for $i = 1, \ldots, 2^d$, be the decomposition of Λ_m into disjoint boxes all of which are translates of Λ_{m-1}. We have

$$\Pi_{m,n}(t) \leqq -\sum_{i=1}^{2^d} T \int q_{\Lambda_m, n}(t)^{-1} \sum_{x \in (\Lambda_{m-1}^i)_r^o} |\nabla_x q_{\Lambda_m, n}(t) + \beta q_{\Lambda_m, n}(t) \nabla_x H|^2 d\sigma. \tag{16}$$

We employ the following inequality, valid for measurable functions f and $g > 0$ on a measure space (X, Σ, μ) by Hölder's inequality:

$$\int g^{-1} f^2 d\mu \geqq (\int f d\mu)^2 (\int g d\mu)^{-1}. \tag{17}$$

Integrating in the ith term on the right side of (16) first over the variables $\xi(x)$ for x in the complement of Λ_{m-1}^i and using inequality (17),

$$\Pi_{m,n}(t) \leqq \sum_{i=1}^{2^d} \Pi_{m-1,n}^i(t),$$

where $\Pi^i_{m-1,n}(t)$ is defined similarly to $\Pi_{m-1,n}(t)$ but with Λ^i_{m-1} replacing Λ_{m-1}. Since $S_n(t)$ commutes with translations $(\mathrm{mod}-2n)$ of Γ_n and v is translation invariant it is easily shown that

$$q_{\Lambda+x,n}(t) = \tau_x q_{\Lambda,n}(t),$$

provided that both Λ and $\Lambda+x$ are in Γ_n. Applying this fact to the integral defining $\Pi^i_{m-1,n}(t)$, we conclude that $\Pi^i_{m-1,n}(t) = \Pi_{m-1,n}(t)$ for each i.

Step 8. Applying Fatou's inequality again we obtain

$$f(v(t)) - f(v) \leqq \int_0^t \Pi(\tau)d\tau, \tag{18}$$

where

$$\Pi(t) = \lim_{m\to\infty} |\Lambda_m|^{-1} \left(\limsup_{n\to\infty} \left\{ -\sum_{x\in(\Lambda_m)^o} T\int q_{\Lambda_m,n}(t)^{-1} |\nabla_x q_{\Lambda_m,n}(t) + \beta q_{\Lambda_m,n}(t)\nabla_x H|^2 d\sigma \right\} \right). \tag{19}$$

This is the required expression.

IV. Proof of the Theorems in Sect. I

The proof of the first claim in Theorem I.1 was completed in Step 6. For the proof of the second statement we need the following proposition and Lemma.

Proposition 2. *Let H have potentials of range r, and let μ be a state on Ξ. Then μ is a Gibbs state of H if and only if it is locally absolutely continuous with respect to product measure σ and for every finite $\Lambda \subset L$ the Radon-Nikodym derivative $q_\Lambda = d(\mu|\Lambda)/d\sigma$ satisfies*

$$\nabla_x q_\Lambda + \beta q_\Lambda \nabla_x H = 0 \tag{1}$$

for all $x\in\Lambda^o_r$.

Remarks. (1) This proposition is closely connected with the characterization of the Gibbs states by their quasi-invariance under rotation of finitely-many spins. (2) For plane-rotor models $(M = S^1)$ it is easily seen that (1) is equivilent to the equations (in the notation of the appendix to [5]):

$$n(x)\hat{\mu}(\mathbf{n}) + \beta \sum_{\mathbf{m}} m(x)\hat{H}(\mathbf{m})\hat{\mu}(\mathbf{n} - \mathbf{m}) = 0$$

holding for all multiindices $\mathbf{n}$, where $\hat{\mu}(\mathbf{n})$ is the Fourier coefficient

$$\hat{\mu}(\mathbf{n}) = \int \bar{\gamma}_\mathbf{n} d\mu.$$

Proof of Proposition 2. Let μ be Gibbs and $\Lambda \subset L$ finite. Let $\varrho_{\Lambda,\zeta}(\cdot)$ be the density with respect to σ of the conditional Gibbs state $\mu_{\Lambda,\zeta}$ (with boundary condition $\zeta\in M^{L\backslash\Lambda}$) defined in [5, Sect. V]. $\varrho_{\Lambda,\zeta}$ is the unique solution of the equations

$$\nabla_x \varrho_{\Lambda,\zeta} + \beta \varrho_{\Lambda,\zeta} \nabla_x H_{\Lambda,\zeta} = 0$$

for $x \in \Lambda$. Let u be any smooth one-form on M^Λ. By pulling back u along the canonical projection: $M^L \to M^\Lambda$, we may regard u as a one-form on Ξ. Using the divergence theorem we have, for all boundary conditions ζ,

$$0 = \int (\nabla \varrho_{\Lambda,\zeta} + \beta \varrho_{\Lambda,\zeta} \nabla H_{\Lambda,\zeta}) \cdot u \, d\sigma$$
$$= \int \varrho_{\Lambda,\zeta} (-\nabla \cdot u + \beta \nabla H_{\Lambda,\zeta} \cdot u) d\sigma .$$

Integrating over ζ with μ and using the DLR equations we obtain

$$0 = \int (-\nabla \cdot u + \beta \nabla H \cdot u) d\mu . \tag{2}$$

Now assume that $u_x = \pi_x u$ is non-zero only if $x \in \Lambda_r^o$. Then $\nabla H \cdot u = \sum_{x \in \Lambda_r^o} \nabla_x H \cdot u_x$

is independent of ζ and so (2) may be written

$$0 = \int q_\Lambda (-\nabla \cdot u + \beta \nabla H \cdot u) d\sigma$$
$$= \int (\nabla q_\Lambda + \beta q_\Lambda \nabla H) \cdot u \, d\sigma .$$

Since u was arbitrary we obtain (1).

Conversely assume that (1) is satisfied. Let $u = g v$ with g a smooth function of $\eta(x)$ for $x \in \partial_r \Lambda_r^o$ and v the pull-back of a one-form on $M^{\Lambda_r^o}$. Then (2) holds for this u. If $\mu(\cdot | \partial_r \Lambda_r^o)(\eta)$ is a regular conditional expectation for μ given the configuration η on $\partial_r \Lambda_r^o$ we obtain from (2):

$$0 = \int d\mu(\eta) g(\eta) \int d\mu(\xi | \partial_r \Lambda_r^o)(\eta) (-\nabla \cdot v(\xi) + \beta v(\xi) \cdot \nabla H_{\Lambda_r^o,\eta}(\xi)) .$$

Letting g and v range over suitable countable dense subsets we conclude that $\mu(\cdot | \partial_r \Lambda_r^o)(\eta)$ is, for μ-a.e. η, a weak solution of the first order system

$$\nabla_x \varrho_\eta + \beta \varrho_\eta \nabla_x H_{\Lambda_r^o,\eta} = 0$$

for $x \in \Lambda_r^o$. Uniqueness and regularity theorems for first order linear PDE.'s then give

$$\varrho_\eta(\cdot) = d\mu(\cdot | \partial_r \Lambda_r^o)(\eta)/d\sigma = \varrho_{\Lambda_r^o,\eta}(\cdot) ,$$

for μ-a.e. η. Thus μ satisfies the DLR equations with Λ_r^o replacing Λ. $\quad\square$

Lemma. *Suppose that μ is not a Gibbs state of the Hamiltonian H. Then there exist positive numbers ε and δ and a neighborhood U of μ in the weak topology such that if v is locally smooth, translation invariant and contained in U then*

$$f(v(t)) - f(v) \leqq -t\delta$$

for $0 \leqq t \leqq \varepsilon$.

Proof of the Lemma. If μ is not Gibbs, then by the preceding proposition there exists a smooth one-form u (the pull-back of a one-form on M^Λ for some finite Λ) and $\delta_1 > 0$ such that

$$T\left(\int (-\nabla \cdot u + \beta \nabla H \cdot u) d\mu\right)^2 > \delta_1 \int |u|^2 d\mu . \tag{3}$$

Define U_1 to be the set of states for which (3) holds. U_1 is open in the weak topology and contains μ. By the continuity of the map $(v,t) \to v(t)$ we can find $\varepsilon > 0$ and a possibly smaller neighborhood U of μ such that $v(t) \in U_1$ if $v \in U$ and $0 \leq t \leq \varepsilon$.

Now let v be locally smooth, translation invariant and contained in U. Defining $q_{A,n}(t)$ as in Step three we obtain (by an application of Hölder's inequality):

$$|\int u \cdot (\nabla q_{A,n}(t) + \beta q_{A,n}(t)\nabla H)d\sigma|$$
$$\leq (\int |u|^2 q_{A,n}(t)d\sigma)^{1/2}(\int q_{A,n}(t)^{-1}|\nabla q_{A,n}(t) + \beta q_{A,n}(t)\nabla H|^2 d\sigma)^{1/2}.$$

After an application of the divergence theorem and making several obvious substitutions we obtain

$$-\sum_{x \in A} T\int q_{A,n}(t)^{-1}|\nabla_x q_{A,n}(t) + \beta q_{A,n}(t)\nabla_x H|^2 d\sigma$$
$$\leq -T(\int(-\nabla \cdot u + \beta \nabla H \cdot u)dv_n(t))^2/(\int |u|^2 dv_n(t)). \tag{4}$$

Since $v_n(t) \to v(t)$ as $n \to \infty$, $v_n(t)$ is eventually in U. Let m_0 be so large that $A \subset (A_{m_0})_r^o$. We conclude from Sect. III, (13), (14) and Sect. IV, (4) that

$$\Pi_{m_0}(t) \leq -\delta_1$$

for $0 \leq t \leq \varepsilon$. Defining $\delta = \delta_1|A_{m_0}|^{-1}$ and remembering that $|A_m|^{-1}\Pi_m(t)$ decreases as $m \to \infty$ we see from Sect. III, (19) that

$$\Pi(t) \leq -\delta$$

for $0 \leq t \leq \varepsilon$. The lemma now follows from the formula of Step eight. $\quad\square$

Returning to the proof of the theorems, we note that the second assertion of Theorem I.1 is an immediate consequence of the lemma. Theorem I.2 follows immediately from the lemma and the observation that the free energy functional is bounded below. [The entropy makes a non-positive contribution to the free energy, and there exists a constant $C < +\infty$ with $H_A(\cdot) \geq -C|A|$.] $\quad\square$

References

1. Dobrushin, R.L., Shlosman, S.B.: Commun. Math. Phys. **42**, 31 (1975)
2. Fröhlich, J., Simon, B., Spencer, T.: Commun. Math. Phys. **50**, 79 (1976)
3. Faris, W.H.: J. Funct. Anal. **32**, 342 (1979)
4. Faris, W.H.: Trans. Am. Math. Soc. **261**, 579 (1980)
5. Wick, W.D.: Commun. Math. Phys. **81**, 361–377 (1981)
6. Ito, K.: The Brownian motion and tensor fields on Riemannian manifolds. Proc. of the International Congress of Mathematicians (1962), pp. 537–539
7. Ruelle, D.: Statistical mechanics. Rigorous results. New York: W.A. Benjamin Inc. 1969
8. Holley, R.: Free energy in Markovian model of a lattice spin system. Commun. Math. Phys. **23**, 87–99 (1971)
9. Higuchi, Y., Shiga, T.: Some results on Markov processes of infinite lattice spin systems. J. Math. Kyoto Univ. **15**, 211–229 (1975)

10. Sullivan, W.G.: Specific information gain for interacting Markov processes. Z. Wahrsch. Verw. Gebiete **37**, 77–90 (1976)
11. Moulin Ollagnier, J., Pinchon, D.: Free energy in spin-flip processes is non-increasing. Commun. Math. Phys. **55**, 29–35 (1977)
12. Holley, R.A., Stroock, D.W.: In one and two dimensions, every stationary measure for a stochastic Ising model is a Gibbs state. Commun. Math. Phys. **55**, 37–45 (1977)
13. Kurtz, T.: Extensions of Trotter's operator semigroup approximation theorems. J. Funct. Anal. **3**, 354–375 (1969)
14. Holley, R., Stroock, D.: Diffusions on an infinite dimensional torus. Preprint (1981)