Existence of Phase Transitions
for Long-Range Interactions*

Robert B. Israel

Princeton University, Princeton, N. J., USA

Abstract. Using a theorem about tangent functionals to convex functions, we obtain existence results for phase transitions. In a "large" Banach space of interactions very pathological behavior is found. In spaces of more "reasonable" interactions we obtain co-existing phases which differ in the expectation of a given observable, as well as broken translation invariance due to long-range order. As an example we consider the isotropic Heisenberg model.

1. Introduction

A standard type of problem in statistical mechanics is to describe the transla-tion-invariant equilibrium states for a given interaction. Here we consider the reverse situation: given an invariant state ϱ, we look for an interaction having an invariant equilibrium state $\tilde{\varrho}$ which bears some resemblance to ϱ. In particular, if ϱ exhibits some type of long-range order, we would like $\tilde{\varrho}$ to share this property. This approach leads to some very general existence results for phase transitions, showing that a given type of phase transition occurs for some member of a certain class of interactions.

Our basic point of view will be the identification of invariant equilibrium states with tangent functionals to the pressure on a Banach space of interactions. This identification can be made for classical and quantum lattice and hard-core continuous systems [6–8], but for the sake of simplicity we will mainly consider quantum lattice systems. The main tool, presented in Section 2, is a theorem on approximation of functionals by tangent functionals to a convex function. In Section 3 we consider a "large" Banach space of interactions, finding very patholo-gical behavior. In Section 4 we deal with more "reasonable" interactions, such as pair interactions, finding phase transitions involving the expectation of a given observable, as well as broken translation invariance due to long-range order. Section 5 applies these methods to an important example, the isotropic Heisenberg model, and Section 6 briefly sketches applications to classical systems.

2. Approximation by Tangent Functionals

Let P be a continuous convex function on a real Banach space $\mathscr{X}$. A linear functional $\alpha \in \mathscr{X}^*$ will be called *P-bounded* if for some constant C, $P(\Psi) \geq \alpha(\Psi) + C$ for all $\Psi \in \mathscr{X}$. We say that α is *tangent to* P at $\Phi \in \mathscr{X}$ if $P(\Phi + \Psi) \geq P(\Phi) + \alpha(\Psi)$ for
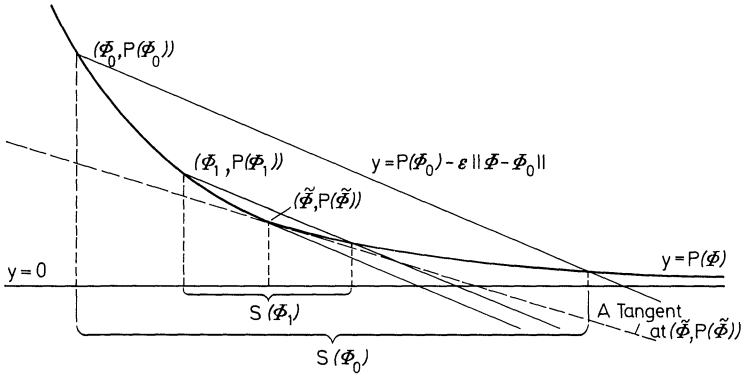

Fig. 1

all $\Psi \in \mathscr{X}$. A theorem of Bishop and Phelps [1] (who consider convex sets rather than convex functions) shows that tangent functionals to P are dense in the P-bounded functionals (with the norm topology). The generalized version presented below allows stronger control over the point of tangency if less control over the tangent functional is demanded.

Theorem 1. *Let $\mathscr{F}$ be a closed convex cone in $\mathscr{X}$ (i.e. if $\Psi_1, \Psi_2 \in \mathscr{F}$ and $c \geq 0$, then $c(\Psi_1 + \Psi_2) \in \mathscr{F}$). Then given $\Phi_0 \in \mathscr{X}, \alpha_0 \in \mathscr{X}^*$ P-bounded, and $\varepsilon > 0$, there is $\tilde{\alpha} \in \mathscr{X}^*$ which is tangent to P at some $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$, with*

$$\tilde{\alpha}(\Psi) \geq \alpha_0(\Psi) - \varepsilon \|\Psi\| \quad for \quad \Psi \in \mathscr{F} \tag{1}$$

and

$$\|\Phi_0 - \tilde{\Phi}\| \leq \varepsilon^{-1}(P(\Phi_0) - \alpha_0(\Phi_0) - C) \tag{2}$$

where $C = \inf\{P(\Psi) - \alpha_0(\Psi) : \Psi \in \mathscr{X}\}$.

Note that if $\mathscr{F}$ is a closed linear subspace of $\mathscr{X}$ we obtain $\|(\alpha_0 - \tilde{\alpha})|\mathscr{F}\| < \varepsilon$.

Proof. By considering the function $P - \alpha_0 - C$ instead of P, we can reduce the theorem to the case where $\alpha_0 = 0$ and $C = 0$. Then we must find $\tilde{\alpha}$ tangent to P at $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$ with $\tilde{\alpha}(\Psi) \geq -\varepsilon \|\Psi\|$ for $\Psi \in \mathscr{F}$ and $\|\Phi_0 - \tilde{\Phi}\| \leq \varepsilon^{-1} P(\Phi_0)$. See Fig. 1.

For each $\Phi \in \mathscr{X}$ we define

$$S(\Phi) = \{\Psi \in \Phi + \mathscr{F} : P(\Psi) \leq P(\Phi) - \varepsilon \|\Phi - \Psi\|\}.$$

We will find $\tilde{\Phi}$ so that $S(\tilde{\Phi}) = \{\tilde{\Phi}\}$. Note that $S(\Phi)$ is closed, and that if $\Psi \in S(\Phi)$ then $S(\Psi) \subset S(\Phi)$. We construct a sequence $\Phi_0, \Phi_1, \ldots$ by choosing $\Phi_{n+1} \in S(\Phi_n)$ with

$$P(\Phi_{n+1}) < 2^{-n}\varepsilon + \inf\{P(\Psi) : \Psi \in S(\Phi_n)\}.$$

Then if $\Psi \in S(\Phi_n)$ with $n \geq 1$, by the choice of Φ_n in $S(\Phi_{n-1})$

$$P(\Phi_n) - 2^{1-n}\varepsilon < P(\Psi) \leq P(\Phi_n) - \varepsilon \|\Psi - \Phi_n\|$$

so $\|\Psi - \Phi_n\| < 2^{1-n}$. Thus the Φ_n form a Cauchy sequence with limit $\tilde{\Phi} \in \bigcap_n S(\Phi_n)$. If $\Psi \in S(\tilde{\Phi}) \subset \bigcap_n S(\Phi_n)$, then $\|\Psi - \Phi_n\| < 2^{1-n}$ for all $n \geq 1$, so $\Psi = \tilde{\Phi}$, i.e. $S(\tilde{\Phi}) = \{\tilde{\Phi}\}$.

Now the open "supergraph" $\{(\Psi, y) \in \mathscr{X} \oplus \mathbb{R} : y > P(\Psi)\}$ and the set $\{(\Psi, y) : \Psi \in \tilde{\Phi} + \mathscr{F}, y \leq P(\tilde{\Phi}) - \varepsilon \|\Psi - \tilde{\Phi}\|\}$ are convex and disjoint, so by the

Hahn-Banach theorem they can be separated by a hyperplane given by a nonzero continuous linear functional on $\mathscr{X} \oplus \mathbb{R}$. This functional can be written as $(\Psi, y) \mapsto y - \tilde{\alpha}(\Psi)$, so that for some $t \in \mathbb{R}$

$$P(\tilde{\Phi}) - \varepsilon \|\Psi - \tilde{\Phi}\| - \tilde{\alpha}(\Psi) \leq t \quad \text{for all} \quad \Psi \in \tilde{\Phi} + \mathscr{F}$$

$$t \leq P(\Psi) - \tilde{\alpha}(\Psi) \quad \text{for all} \quad \Psi \in \mathscr{X} .$$

Taking in particular $\Psi = \tilde{\Phi}$ we must have $t = P(\tilde{\Phi}) - \tilde{\alpha}(\tilde{\Phi})$. Thus the inequalities become

$$\tilde{\alpha}(\Psi - \tilde{\Phi}) \geq -\varepsilon \|\Psi - \tilde{\Phi}\| \quad \text{for} \quad \Psi \in \tilde{\Phi} + \mathscr{F}$$

$$P(\Psi) \geq P(\tilde{\Phi}) + \tilde{\alpha}(\Psi - \tilde{\Phi}) \quad \text{for} \quad \Psi \in \mathscr{X} .$$

This proves the theorem.

3. The Space $\mathscr{B}$

Here is a brief summary of the mathematical framework for quantum lattice systems (see [6]).

1) At each point i of the lattice $\mathbb{Z}^\nu$ there is a copy $\mathscr{H}_i$ of a finite-dimensional Hilbert space $\mathscr{H}$. For each finite subset X of $\mathbb{Z}^\nu$ there is the Hilbert space $\mathscr{H}_X = \bigotimes_{i \in X} \mathscr{H}_i$.

2) The *quasi-local C* algebra* $\mathfrak{A}$ is the inductive limit of the local algebras $\mathfrak{A}_X$ of all linear operators on $\mathscr{H}_X$, with the natural injections. Translations in the lattice act on $\mathfrak{A}$ by automorphisms τ_i, $i \in \mathbb{Z}^\nu$. The set of translation-invariant states of $\mathfrak{A}$ is denoted E^I.

3) An *interaction* Φ is a function from finite nonempty subsets X of $\mathbb{Z}^\nu$ to self-adjoint operators $\Phi(X) \in \mathfrak{A}_X$, with
   (i) $\Phi(i + X) = \tau_i \Phi(X)$ for $i \in \mathbb{Z}^\nu$,
   (ii) $\|\Phi\| \equiv \sum_{X \ni 0} N(X)^{-1} \|\Phi(X)\| < \infty$,
where $N(X)$ is the number of points in X. The interactions form a real separable Banach space $\mathscr{B}$ in the above norm.

For $\Phi \in \mathscr{B}$ we define $A_\Phi = \sum_{X \ni 0} N(X)^{-1} \Phi(X) \in \mathfrak{A}$.

4) There is an affine upper semicontinuous function s on E^I (with the weak-* topology), called the *mean entropy*.

5) There is a convex function P on $\mathscr{B}$, called the *pressure*, with $|P(\Phi) - P(\Psi)| \leq \|\Phi - \Psi\|$ for $\Phi, \Psi \in \mathscr{B}$.

6) For any $\Phi \in \mathscr{B}$, $P(\Phi) = \sup \{s(\varrho) - \varrho(A_\Phi) : \varrho \in E^I\}$. If $P(\Phi) = s(\varrho) - \varrho(A_\Phi)$, we say ϱ is an *invariant equilibrium state* for Φ.

7) For any $\varrho \in E^I$, $s(\varrho) = \inf \{P(\Phi) + \varrho(A_\Phi) : \Phi \in \mathscr{B}\}$. Thus the functional α on $\mathscr{B}$ given by $\alpha(\Phi) = -\varrho(A_\Phi)$ is P-bounded. The relation $\alpha(\Phi) = -\varrho(A_\Phi)$ sets up a $1 - 1$ correspondence between functionals tangent to P at $\Psi \in \mathscr{B}$ and invariant equilibrium states for Ψ.

8) Extreme points of E^I are called *ergodic states*. Any $\varrho \in E^I$ is the barycenter of a unique probability measure μ concentrated on the ergodic states (this is called the *ergodic decomposition* of ϱ). If ϱ is an invariant equilibrium state for Φ, then the support of μ consists of invariant equilibrium states for Φ.

In applying Theorem 1 to quantum lattice systems, we take $\mathscr{X} = \mathscr{B}$ and P the pressure. Note that if $\varrho, \varrho' \in E^I$ with corresponding functionals α, α' on $\mathscr{B}$ given as above, then $\|\varrho - \varrho'\| = \|\alpha - \alpha'\|$. Since $\|A_\Phi\| \leq \|\Phi\|$, we clearly have $\|\varrho - \varrho'\| \geq \|\alpha - \alpha'\|$. Conversely, for $B \in \mathfrak{A}_X$ selfadjoint, there is the interaction Ψ_B defined by

$$\Psi_B(i + X) = \tau_i B \quad \text{for} \quad i \in \mathbb{Z}^\nu$$
$$\Psi_B(Y) = 0 \text{ if } Y \text{ is not a translate of } X . \tag{3}$$

Since $\|\Psi_B\| = \|B\|$ and $\varrho(A_{\Psi_B}) = \varrho(B)$ for all $\varrho \in E^I$, we have $|\varrho(B) - \varrho'(B)| \leq \|\alpha - \alpha'\| \, \|B\|$. It is then easily seen that $\|\varrho - \varrho'\| = \|\alpha - \alpha'\|$ as claimed. Taking $\mathscr{F}$ as the whole space $\mathscr{B}$, Theorem 1 then asserts the following:

Corollary 1. *For any $\Phi_0 \in \mathscr{B}$, $\varrho_0 \in E^I$, and $\varepsilon > 0$, there is an invariant equilibrium state $\tilde{\varrho}$ for some $\tilde{\Phi} \in \mathscr{B}$ with*

$$\|\varrho_0 - \tilde{\varrho}\| \leq \varepsilon , \tag{4}$$

and

$$\|\Phi_0 - \tilde{\Phi}\| \leq \varepsilon^{-1} \left(P(\Phi_0) + \varrho_0(A_{\Phi_0}) - s(\varrho_0) \right). \tag{5}$$

The norm condition (4) gives us very strong control over $\tilde{\varrho}$. This can best be seen by considering the ergodic decomposition.

Theorem 2. *Let $\mathfrak{A}$ be any separable G-abelian C^* algebra with identity (see [10], Chapter 6). For an invariant state ϱ let μ_ϱ be the unique probability measure with barycenter ϱ concentrated on the ergodic states of $\mathfrak{A}$. Then for any invariant states ϱ, ϱ' on $\mathfrak{A}$, $\|\varrho - \varrho'\| = \|\mu_\varrho - \mu_{\varrho'}\|$.*

Proof. It is clear that $\|\varrho - \varrho'\| \leq \|\mu_\varrho - \mu_{\varrho'}\|$. By a theorem of Grothendieck [5] any hermitian linear functional h on $\mathfrak{A}$ has a unique decomposition $h = h_+ - h_-$ with h_+ and h_- positive and $\|h\| = \|h_+\| + \|h_-\|$. We take $h = \varrho - \varrho'$. Since h is G-invariant and the decomposition fs unique, h_+ and h_- are also invariant. Thus h_+ and h_- are represented by positive measures μ_+ and μ_- concentrated on the ergodic states of $\mathfrak{A}$, with $\|\mu_+\| = \|h_+\|$ and $\|\mu_-\| = \|h_-\|$. Then since $\varrho + h_- = \varrho' + h_+$ and the ergodic decomposition is unique, we must have $\mu_\varrho + \mu_- = \mu_{\varrho'} + \mu_+$. Therefore

$$\|\varrho - \varrho'\| = \|h_+\| + \|h_-\| = \|\mu_+\| + \|\mu_-\|$$
$$\geq \|\mu_+ - \mu_-\| = \|\mu_\varrho - \mu_{\varrho'}\|$$

which completes the proof.

Theorem 3. a) *If $\varrho_1, \ldots, \varrho_n$ are ergodic states, there is some interaction $\tilde{\Phi} \in \mathscr{B}$ for which these are all invariant equilibrium states.*

b) *If μ is a nonatomic probability measure concentrated on the ergodic states, there is some $\tilde{\Phi} \in \mathscr{B}$ which has uncountably many ergodic equilibrium states in the support of μ.*

c) *There is a dense set of interactions with uncountably many ergodic equilibrium states.*

Proof. For (a) let $\varrho_0 = n^{-1}(\varrho_1 + \cdots + \varrho_n)$ and take $\varepsilon < n^{-1}$. If $\tilde{\varrho}$ is an invariant equilibrium state for $\tilde{\Phi}$ and $\|\varrho_0 - \tilde{\varrho}\| \leq \varepsilon$, then in the ergodic decomposition $\mu_{\tilde{\varrho}}(\{\varrho_i\}) > 0$ for each $i = 1, \ldots, n$, and so the ϱ_i are all invariant equilibrium states for $\tilde{\Phi}$. In (b) let ϱ_0 be the barycenter of μ, and let $\tilde{\varrho}$ be an invariant equilibrium state

for $\tilde{\Phi}$ with $\|\varrho_0 - \tilde{\varrho}\| < 2$. Then $\|\mu - \mu_{\tilde{\varrho}}\| < 2$, and there must be uncountably many ergodic states in the intersection of the supports of μ and $\mu_{\tilde{\varrho}}$, all of which are ergodic equilibrium states for $\tilde{\Phi}$. This method and the estimate (5) will also give us (c) if we can show that for any $\Phi_0 \in \mathcal{B}$ and $\delta > 0$ there is a nonatomic probability measure μ concentrated on ergodic states ϱ with $P(\Phi_0) + \varrho(A_{\Phi_0}) - s(\varrho) \leq \delta$. It can be shown that ergodic states form a dense G_δ set in the compact metrizable space $\{\varrho \in E^I : P(\Phi_0) + \varrho(A_{\Phi_0}) - s(\varrho) \leq \delta\}$ (with the weak-* topology; metrizability follows from separability of $\mathfrak{A}$). The measure μ is then obtained by constructing a homeomorphic image of the Cantor set in this G_δ, and taking a nonatomic probability measure on the Cantor set. (See [13] for further details.)

It is shown in [9] that in a certain dense subspace of $\mathcal{B}$ any two interactions which share an equilibrium state must be "physically equivalent" modulo a constant term. By Theorem 3, this result fails in $\mathcal{B}$ itself in the most spectacular manner imaginable: the set of interactions having a given ergodic equilibrium state is so large and diverse that any other ergodic state is an invariant equilibrium state for some interaction in that set. Fisher [3] has obtained examples of classical systems where two different interactions share an invariant equilibrium state, but the behavior described by Theorem 3 is much wilder than that of his models. Remarkable as this behavior is, it does not appear to be at all related to the phase transitions of more "reasonable" systems, and should be considered as a pathology arising from the use of too large a space of interactions. If we use a norm which gives a higher weighting to terms $\Phi(X)$ with $N(X)$ large, we will no longer be able to obtain such pathologies.

4. Phase Transitions and Cluster Properties

In this section we will apply Theorem 1 to quantum lattice systems with $\mathcal{F}$ containing a much more restricted class of interactions. If $B \in \mathfrak{A}_X$ is self-adjoint, this will yield an estimate $\tilde{\varrho}(B) \leq \varrho_0(B) + \varepsilon \|B\|$ if $\Psi_B \in \mathcal{F}$, where Ψ_B was defined in (3). To obtain non-ergodic invariant equilibrium states or breaking of translation invariances from this type of information, we use cluster properties of states.

For $A \in \mathfrak{A}$ let $S_n(A) = n^{-2\nu} \left(\sum_{i \in C_n} \tau_i A \right)^2$ where C_n is a cube of side n in $\mathbb{Z}^\nu$. Ergodic states are characterized by the following "weak cluster property" ([10], Lemma 6.5.1): $\lim_{n \to \infty} \varrho(S_n(A)) = \varrho(A)^2$ for all $A \in \mathfrak{A}$ self-adjoint. The above limit will exist for all invariant states ϱ and is always greater than or equal to $\varrho(A)^2$; this is easily seen using the Gel'fand-Naimark-Segal representation and the "mean ergodic theorem" ([10], Proposition 6.2.15). If $\varrho \in E^I$ with ergodic decomposition $\varrho = \int \varrho' \mu(d\varrho')$ and $\lim_{n \to \infty} \varrho(S_n(A)) \neq \varrho(A)^2$, then by the Lebesgue Dominated Convergence Theorem

$$\int \varrho'(A)^2 \mu(d\varrho') = \lim_{n \to \infty} \int \varrho'(S_n(A)) \mu(d\varrho')$$
$$= \lim_{n \to \infty} \varrho(S_n(A)) \neq \varrho(A)^2.$$

Thus $\varrho'(A) \neq \varrho(A)$ for some ϱ' in the support of μ. This means that control over expectations of A and of $A\tau_i A$ for all $i \in \mathbb{Z}^\nu$ (or at least sufficiently much of $\mathbb{Z}^\nu$) can lead to the existence of a phase transition.

Theorem 4. *Let $A \in \mathfrak{A}_X$ be self-adjoint, and suppose there are invariant states that differ on A. Let $\mathscr{F}$ be the cone of interactions Ψ of the form*

$$\Psi(i + X) = h\tau_i A, \qquad h \in \mathbb{R},$$

$$\Psi((i + X) \cup (j + X)) = J(i - j)(\tau_i A)(\tau_j A) \quad if \quad (i + X) \cap (j + X) = \emptyset,$$

$$with \quad J(i - j) \leq 0,$$

$$\Psi(Y) = 0 \quad for \ all \ other \ Y.$$

Then for any $\Phi_0 \in \mathscr{B}$ there is $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$ which has at least two invariant equilibrium states that differ on A.

Proof. Let $\varrho_0 = (\varrho_1 + \varrho_2)/2$, where $\varrho_1, \varrho_2 \in E^I$ differ on A. Then $\lim_{n \to \infty} \varrho_0(S_n(A)) \geq (\varrho_1(A)^2 + \varrho_2(A)^2)/2 > \varrho_0(A)^2$. By Theorem 1, for any $\varepsilon > 0$ there is an invariant equilibrium state $\tilde{\varrho}$ for some $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$ with $|\tilde{\varrho}(A) - \varrho_0(A)| \leq \varepsilon \|A\|$ and $\tilde{\varrho}(A\tau_i A) \geq \varrho_0(A\tau_i A) - \varepsilon \|A\|^2$ when $(i + X) \cap X = \emptyset$. Thus

$$\lim_{n \to \infty} \tilde{\varrho}(S_n(A)) \geq \lim_{n \to \infty} \varrho_0(S_n(A)) - \varepsilon \|A\|^2$$

and

$$\tilde{\varrho}(A)^2 \leq \varrho_0(A)^2 + 2\varepsilon \|A\|^2.$$

So for ε sufficiently small, $\tilde{\varrho}$ decomposes into ergodic equilibrium states for $\tilde{\Phi}$ which differ on A.

To deal with breaking of translation invariance we use the following property, equivalent to the state ϱ having *short range correlations* [7, 11]:

For any $\varepsilon > 0$ and $A \in \mathfrak{A}$ there is a finite subset X of $\mathbb{Z}^\nu$ such that if $B \in \mathfrak{A}_Y$ with $X \cap Y = \emptyset$, then

$$|\varrho(AB) - \varrho(A)\varrho(B)| \leq \varepsilon \|B\|.$$

The *decomposition at infinity*, a finer decomposition than the ergodic decomposition, represents any state as the barycenter of a probability measure carried by states with short range correlations. A KMS state for some interaction is decomposed in this way into extremal KMS states. Thus, given an invariant equilibrium state for some interaction, we can first decompose it into ergodic equilibrium states, and then further decompose any of these that lack short range correlations into non-invariant states with short range correlations; the latter will be KMS states if the interaction is in the space $\mathscr{B}_1$ on which time evolution is defined ([6], Theorem 4). If some ergodic equilibrium state for an interaction lacks short range correlations, we say that translation invariance is broken.

Theorem 5. *Let $A \in \mathfrak{A}_X$ and $B \in \mathfrak{A}_Y$ be self-adjoint, and suppose $\lim_{|i| \to \infty} \varrho_0(A\tau_i B)$ does not exist for some invariant state ϱ_0. Let $\mathscr{F}$ be the closed linear span of the interactions $\Psi_{A\tau_i B}$ (defined as in (3)) for $X \cap (i + Y) = \emptyset$. Then for any $\Phi_0 \in \mathscr{B}$ there is $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$ which has some ergodic equilibrium state ϱ' such that $\lim_{|i| \to \infty} \varrho'(A\tau_i A)$ does not exist. Thus ϱ' does not have short range correlations, and translation invariance is broken for $\tilde{\Phi}$.*

Proof. By Theorem 1, with $\varepsilon > 0$ sufficiently small, there is an invariant equilibrium state $\tilde{\varrho}$ for some $\tilde{\Phi} \in \Phi_0 + \mathscr{F}$ such that $\tilde{\varrho}(A\tau_i B)$ does not tend to a

limit. By the Lebesgue Dominated Convergence Theorem, $\lim_{|i| \to \infty} \varrho'(A\tau_i B)$ fails to exist for some ergodic state ϱ' in the ergodic decomposition of $\tilde{\varrho}$. But if ϱ' had short range correlations, $\varrho'(A\tau_i B)$ would tend to $\varrho'(A)\varrho'(B)$ as $|i| \to \infty$.

5. The Isotropic Heisenberg Model

The Ising and Heisenberg models are important examples to which we can apply the methods we have developed. For the Ising model a good deal is known: in one dimension there is spontaneous magnetization for ferromagnetic interactions decaying as $|i-j|^{-r}$ with $1 < r < 2$ [2], and in two or more dimensions for nearest-neighbor interactions. In contrast, there are no (translation-invariant) isotropic Heisenberg interactions in any dimension for which a phase transition has been proven to exist. Thus our existence results will be especially interesting for these interactions. A very similar analysis will apply to the Ising model.

We take the Hilbert space $\mathscr{H}_i$ at each site $i \in \mathbb{Z}^\nu$ to be two-dimensional, with the usual Pauli matrices σ_i^x, σ_i^y, and σ_i^z. Let $\mathscr{Y}_H$ be the space of interactions Φ_J with

$$\Phi_J\{i,j\} = J(i-j)\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \quad \text{for} \quad i \neq j \in \mathbb{Z}^\nu$$

$$\Phi_J(X) = 0 \quad \text{if} \quad N(X) \neq 2 .$$

Since $\|\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j\| = 3$ we have $\|\Phi_J\| = \frac{3}{2} \sum_{i \neq 0} |J(i)| < \infty$. In addition there are one-body interactions ("external fields") given by $\Phi_h\{i\} = \boldsymbol{h} \cdot \boldsymbol{\sigma}_i$ with $\boldsymbol{h} \in \mathbb{R}^3$, and $\|\Phi_h\| = |\boldsymbol{h}|$ (the Euclidean norm).

The "ferromagnetic" phase transition with invariant equilibrium states differing on σ_i can be treated with a slight modification of Theorem 4 (to deal with $\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j$ rather than an ordinary product of operators). Let ϱ_+ and ϱ_- be the states in which all spins are "up" or "down" respectively ($\varrho_\pm(\sigma_i^z) = \pm 1$) and take $\varrho_0 = (\varrho_+ + \varrho_-)/2$. Let $\mathscr{F}$ be the cone of interactions $\Phi_h + \Phi_J$ with $\boldsymbol{h} \in \mathbb{R}^3$, $\Phi_J \in \mathscr{Y}_H$, and $J(i) \leq 0$ for all i. By Theorem 1, for any $\Phi_0 \in \mathscr{B}$ and $\varepsilon > 0$ there is an invariant equilibrium state $\tilde{\varrho}$ for some interaction $\tilde{\Phi} = \Phi_0 + \Phi_h + \Phi_J \in \Phi_0 + \mathscr{F}$ such that $|\tilde{\varrho}(\boldsymbol{u} \cdot \boldsymbol{\sigma}_i)| \leq \varepsilon|\boldsymbol{u}|$ for all $\boldsymbol{u} \in \mathbb{R}^3$, $\tilde{\varrho}(\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j) \geq 1 - 3\varepsilon$ for all $i, j \in \mathbb{Z}^\nu$, and

$$\|\tilde{\Phi} - \Phi_0\| = |\boldsymbol{h}| + \frac{3}{2} \sum_{i \neq 0} |J(i)| \leq \varepsilon^{-1}(P(\Phi_0) + \varrho_0(A_{\Phi_0}))$$

[note $s(\varrho_0) = 0$ since $\varrho_\pm$ are pure states]. We have

$$\tilde{\varrho}(S_n(\sigma_i^x) + S_n(\sigma_i^y) + S_n(\sigma_i^z)) \geq 1 - 3\varepsilon .$$

Thus for $\varepsilon \leq \frac{1}{4}$, $\lim_{n \to \infty} \tilde{\varrho}(S_n(\sigma_i^d)) > \tilde{\varrho}(\sigma_i^d)^2$ where d is one of x, y, and z, so $\tilde{\varrho}$ decomposes into ergodic equilibrium states for $\tilde{\Phi}$ some of which differ on σ_i^d.

If Φ_0 contains only terms which are even polynomials in the spins, we can dispense with the external field: the conditions $\tilde{\varrho}(\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j) \geq 1 - 3\varepsilon$ alone ensure a phase transition at $\tilde{\Phi} = \Phi_0 + \Phi_J$. This is obtained by using the appropriate symmetry, namely time inversion. This is a *-antiautomorphism T of $\mathfrak{A}$ [i.e. an invertible positivity-preserving linear map with $T(AB) = T(B) T(A)$] determined by $T(\boldsymbol{\sigma}_i) = -\boldsymbol{\sigma}_i$. Thus T acts on states by $\varrho \mapsto \varrho \circ T$, and on interactions by $\Phi \mapsto T \circ \Phi$. The pressure is invariant under T, so ϱ is an invariant equilibrium state for Φ if and only if $\varrho \circ T$ is one for $T \circ \Phi$.

Theorem 6. *If $\Phi_0 \in \mathscr{B}$ is invariant under T (meaning each $\Phi_0(X)$ is an even polynomial in the spins σ_i^d, $i \in X$, $d = x, y, z$) there is $\tilde{\Phi} = \Phi_0 + \Phi_J$ with $\Phi_J \in \mathscr{Y}_H$, all $J(i) \leqq 0$, for which there is a phase transition with spontaneous magnetization. That is, for some $u \in \mathbb{R}^3$ and some invariant equilibrium state ϱ' for $\tilde{\Phi}$, $\varrho'(u \cdot \sigma_i) > 0$.*

Proof. Taking $\varepsilon < \frac{1}{3}$ there is $\tilde{\Phi}$ as above with an invariant equilibrium state $\tilde{\varrho}$ for which $\tilde{\varrho}(\sigma_i \cdot \sigma_j) \geqq 1 - 3\varepsilon$ for all $i, j \in \mathbb{Z}^\nu$. Then some ergodic equilibrium state ϱ' in the ergodic decomposition of $\tilde{\varrho}$ satisfies

$$\varrho'(\sigma_i^x)^2 + \varrho'(\sigma_i^y)^2 + \varrho'(\sigma_i^z)^2 = \lim_{n \to \infty} \varrho'\big(S_n(\sigma_i^x) + S_n(\sigma_i^y) + S_n(\sigma_i^z)\big) \geqq 1 - 3\varepsilon > 0.$$

Thus if u is the vector $\langle \varrho'(\sigma_i^x), \varrho'(\sigma_i^y), \varrho'(\sigma_i^z) \rangle$ we have $\varrho'(u \cdot \sigma_i) > 0$. But since $\tilde{\Phi}$ is invariant under T, $\varrho' \circ T$ is also an invariant equilibrium state for $\tilde{\Phi}$, and

$$(\varrho' \circ T)(u \cdot \sigma_i) = -\varrho'(u \cdot \sigma_i) < 0.$$

The bound obtained on $\|\tilde{\Phi}\|$ in the case $\Phi_0 = 0$ using estimate (2) is of some interest, both for Heisenberg and Ising models. For the Ising model, Griffiths [4] has shown using correlation inequalities that there is no spontaneous magnetization if $\sum_{i \neq 0} \tanh |J(i)| < 1$, and hence a fortiori if $\sum_{i \neq 0} |J(i)| < 1$. We will show that this "mean-field" bound is actually the best possible. For the Heisenberg model, Greenberg [12] has shown there is no spontaneous magnetization if $\sum_{i \neq 0} |J(i)| < 0.025$.

Corollary 2. *Let R be the infimum of $\sum_{i \neq 0} |J(i)|$ for interactions $\Phi_J \in \mathscr{Y}_H$ with all $J(i) \leqq 0$, for which there is a spontaneous magnetization, i.e. some invariant equilibrium state ϱ' with $\varrho'(u \cdot \sigma_i) > 0$ for some $u \in \mathbb{R}^3$. Then $R \leqq 1$. The analogous result also holds for the Ising model, where in fact $R = 1$.*

Proof. We need to use a more suitable state than the ϱ_0 above, which would only give $R \leqq 2 \ln 2$. For $m > 0$ let ϱ_m be the state in which spins at different sites are independent, each one being in the ± 1 eigenstate of σ^z with probability $(1 \pm m)/2$; i.e. ϱ_m is the product of states at each site with density matrix

$$\begin{pmatrix} (1+m)/2 & 0 \\ 0 & (1-m)/2 \end{pmatrix}.$$

Then $\varrho_m(\sigma_i \cdot \sigma_j) = m^2$ for $i \neq j$, so a phase transition with spontaneous magnetization is obtained for $\varepsilon < m^2/3$ by the previous methods. The estimate (2) yields

$$R \leqq 2m^{-2}\big(P(0) - s(\varrho_m)\big)$$

where $P(0) = \ln 2$ and

$$s(\varrho_m) = -\frac{1+m}{2} \ln \frac{1+m}{2} - \frac{1-m}{2} \ln \frac{1-m}{2}.$$

As $m \to 0$ we obtain $R \leqq 1$. The same analysis yields $R \leqq 1$ for the Ising model, and by the Griffiths result $R = 1$ there.

Next we examine a case of "antiferromagnetic" breaking of translation invariance in the Heisenberg model. Let A and B be the two sublattices of $\mathbb{Z}^\nu$ consisting of sites i with $|i| \equiv |i_1| + \cdots + |i_\nu|$ even and odd respectively.

Theorem 7. *For any* $\Phi_0 \in \mathcal{B}$ *there is* $\tilde{\Phi} = \Phi_0 + \Phi_J \in \Phi_0 + \mathcal{Y}_H$ *with* $J(i) = 0$ *for* $i \in A$ *and* $J(i) \geqq 0$ *for* $i \in B$, *which exhibits broken translation invariance because* $\sup_{i \in B} \tilde{\varrho}(\sigma_0 \cdot \sigma_i) < 0$ *for some invariant equilibrium state* $\tilde{\varrho}$. *If* $\mathbf{h} \in \mathbb{R}^3$ *let* $R(\mathbf{h})$ *be the infimum of* $\sum_{i \in B} J(i)$ *over interactions* Φ_J *as above for which* $\Phi_{\mathbf{h}} + \Phi_J$ *exhibits this broken translation invariance. Then* $R(\mathbf{h}) \leqq 2 \ln (2 \cosh |\mathbf{h}|)$ *and* $R(0) \leqq 1$.

Proof. Let ϱ_a be the state with $\varrho_a(\sigma_i^z) = (-1)^{|i|}$, i.e. spins in A are all "up" and spins in B are "down". For $j \in B$ the state $\varrho_a \circ \tau_j = \varrho_a \circ T$ has spins in A "down" and spins in B "up". The translation-invariant state $\varrho_0 = (\varrho_a + \varrho_a \circ T)/2$ then has $\varrho_0(\mathbf{u} \cdot \sigma_i) = 0$ and $\varrho_0(\sigma_i \cdot \sigma_j) = (-1)^{|i-j|}$ for $\mathbf{u} \in \mathbb{R}^3$, $i, j \in \mathbb{Z}^\nu$, and $s(\varrho_0) = 0$. By Theorem 1, for any $\Phi_0 \in \mathcal{B}$ and $\varepsilon > 0$ there is an invariant equilibrium state $\tilde{\varrho}$ for some $\tilde{\Phi} = \Phi_0 + \Phi_J \in \Phi_0 + \mathcal{Y}_H$ with $J(i) = 0$ for $i \in A$ and $J(i) \geqq 0$ for $i \in B$, such that $\tilde{\varrho}(\sigma_i \cdot \sigma_j) \leqq -1 + 3\varepsilon$ for $i - j \in B$, and

$$\sum_{i \in B} J(i) \leqq \frac{2}{3\varepsilon} \left(P(\Phi_0) + \varrho_0(A_{\Phi_0}) \right).$$

If $\varepsilon < \frac{1}{3}$ long-range order is obtained, because $\sup_{i \in B} \tilde{\varrho}(\sigma_0 \cdot \sigma_i) < 0$ while since $\lim_{n \to \infty} \tilde{\varrho}(S_n(\sigma_0^x) + S_n(\sigma_0^y) + S_n(\sigma_0^z)) \geqq 0$ we must have $\lim_{|i| \to \infty} \sup \tilde{\varrho}(\sigma_0 \cdot \sigma_i) \geqq 0$. In the case where $\Phi_0 = \Phi_{\mathbf{h}}$ is an external field, $P(\Phi_{\mathbf{h}}) = \ln (2 \cosh |\mathbf{h}|)$ and $\varrho_0(A_{\Phi_{\mathbf{h}}}) = 0$. Thus $R(\mathbf{h}) \leqq 2 \ln (2 \cosh |\mathbf{h}|)$. For $\mathbf{h} = 0$ we can do better by replacing ϱ_a with the state having different lattice sites independent and density matrices $\begin{pmatrix} (1+m)/2 & 0 \\ 0 & (1-m)/2 \end{pmatrix}$ at A sites, $\begin{pmatrix} (1-m)/2 & 0 \\ 0 & (1+m)/2 \end{pmatrix}$ at B sites. Taking $m \to 0$ we obtain $R(0) \leqq 1$, just as in the ferromagnetic case.

6. Classical Systems

The application of our methods to classical lattice systems is fairly straightforward. For the results of Section 4 we may consider a classical lattice-gas or spin system as a special case of the quantum system. However, to obtain results analogous to those of Section 3 we must consider a wider class of interactions than the usual lattice-gas or spin types. For an interaction Φ of this new type, each $\Phi(X)$ will be a real-valued function on the configuration space $\{0, 1\}^X$; thus the Hamiltonian for a finite subset Λ of $\mathbb{Z}^\nu$ will be the function

$$H_\Lambda(\omega) = \sum_{X \subset \Lambda} \Phi(X)(\omega) \qquad \omega \in \{0, 1\}^\Lambda.$$

All our results carry over naturally to such systems, where we replace $\mathfrak{A}$ by the algebra of continuous functions on $\{0, 1\}^{\mathbb{Z}^\nu}$ and states by probability measures on $\{0, 1\}^{\mathbb{Z}^\nu}$. This formalism can also be generalized to deal with higher-spin or continuous-spin classical systems, where $\{0, 1\}$ is replaced by some compact metric space with an "a priori" probability measure. For more details see [13].

Classical hard-core continuous systems, as discussed in [7], also provide a framework for results similar to those of Section 4. The case of two-body interac-

tions is especially interesting. Here a two-body interaction is given by a continuous function V on $\{x \in \mathbb{R}^\nu : |x| \geqq a\}$ where a is the diameter of the hard core, with norm

$$\|\Phi_V\| = \sup_{0 \in X} \sum_{x \in X} |V(x)| < \infty \,,$$

where the supremum is taken over all subsets X of $\mathbb{R}^\nu$ containing 0 such that $|x - y| \geqq a$ for all distinct $x, y \in X$. There are also one-body interactions Φ_μ given by a chemical potential $\mu \in \mathbb{R}$. The methods of Section 4 then lead to the following results:

Theorem 8. a) *For any interaction Φ_0 there is a chemical potential $\mu \in \mathbb{R}$ and a two-body interaction Φ_V with $V \leqq 0$ spherically symmetric, such that $\Phi_0 + \Phi_\mu + \Phi_V$ has two invariant equilibrium states which have different densities.*

b) *For any interaction Φ_0 there is a two-body interaction Φ_V such that translation invariance is broken for $\Phi_0 + \Phi_V$, because some ergodic equilibrium state for $\Phi_0 + \Phi_V$ exhibits long-range order in the two-point correlation function.*

It should be noted that, while part (a) above might be supposed to have some relevance to "liquid-gas" phase transitions, part (b) does not involve a satisfactory kind of "crystallization" in more than one dimension, for which we would want a spherically symmetric interaction and breaking of both rotation and translation invariance.

References

1. Bishop, E., Phelps, R. R.: In: Proc. Sympos. Pure Math., Vol. 7, pp. 27—35. Providence, R.I.: Amer. Math. Soc. 1963
2. Dyson, F.: Commun. math. Phys. **12**, 91—107 (1969)
3. Fisher, M. E.: Commun. math. Phys. **26**, 6—14 (1972)
4. Griffiths, R. B.: Commun. math. Phys. **6**, 121—127 (1967)
5. Grothendieck, A.: J. de Math. pures et appl. **36**, 97—108 (1957)
6. Lanford, O. E. III, Robinson, D. W.: Commun. math. Phys. **9**, 327—338 (1968)
7. Lanford, O. E. III, Ruelle, D.: Commun. math. Phys. **13**, 194—215 (1969)
8. Miracle-Sole, S., Robinson, D. W.: Commun. math. Phys. **19**, 204—218 (1970)
9. Roos, H.: Commun. math. Phys. **36**, 263—276 (1974)
10. Ruelle, D.: Statistical mechanics: rigorous results. New York-Amsterdam: Benjamin 1969
11. Ruelle, D.: In: Kastler, D. (Ed.): Cargese lectures in physics, Vol. 4, pp. 169—194. New York-London-Paris: Gordon and Breach 1970
12. Greenberg, W.: Commun. math. Phys. **13**, 335—344 (1969)
13. Israel, R. B.: Princeton Ph. D. thesis (to appear)

Communicated by G. Gallavotti

R. B. Israel
Department of Mathematics
University of British Columbia
Vancouver, B.C., Canada V6T 1W5