Journal of Statistical Physics, Vol. 52, Nos. 1/2, 1988

Multiple Phase Transitions in the Generalized Curie–Weiss Model

Theodor Eisele[1] **and Richard S. Ellis**[2]

Received December 3, 1987

The generalized Curie–Weiss model is an extension of the classical Curie–Weiss model in which the quadratic interaction function of the mean spin value is replaced by a more general interaction function. It is shown that the generalized Curie–Weiss model can have a sequence of phase transitions at different critical temperatures. Both first-order and second-order phase transitions can occur, and explicit criteria for the two types are given. Three examples of generalized Curie–Weiss models are worked out in detail, including one example with infinitely many phase transitions. A number of results are derived using large-deviation techniques.

KEY WORDS: Generalized Curie–Weiss model; specific Gibbs free energy; large deviations; first-order phase transition; second-order phase transition.

1. INTRODUCTION

The classical Curie–Weiss model is an exactly soluble model of ferromagnetism that allows one to study in detail the behavior of thermodynamic quantities in the neighborhood of the critical point. Unfortunately, the predictions of the model do not completely agree with experiment, and so other models, such as nearest neighbor Ising models, must be considered. However, because of its simplicity and because of the correctness of at least some of its predictions, the classical Curie–Weiss model occupies a central place in the statistical mechanics literature.

The classical Curie–Weiss model is a spin-1/2 model whose Hamiltonian is a quadratic function of the mean spin value in the system.

[1] Institut für Angewandte Mathematik, Universität Heidelberg, D-6900 Heidelberg 1, Federal Republic of Germany.
[2] Department of Mathematics and Statistics, University of Massachusetts, Amherst, Massachusetts 01003.

0022-4715/88/0700-0161$06.00/0

A natural extension of the model is to replace both the quadratic interaction function by a more general interaction function and the spin-1/2 single-site distribution by a more general distribution. The purpose of this paper is to study phase transitions in the resulting models, which are called generalized Curie–Weiss models and are defined below. We will see that the generalized Curie–Weiss model exhibits a number of phenomena that are absent in the classical case.

The generalized Curie–Weiss model is defined in terms of a function g and a probability measure ρ satisfying the following hypotheses.

Hypotheses 1.1. (a) g is an even, real analytic function on $\mathbb{R}$ and is strictly increasing on $[0, \infty)$; $g(0) = 0$.

(b) ρ is a symmetric Borel probability measure on $\mathbb{R}$ that is nondegenerate (i.e., $\rho \neq \delta_0$).

(c) Define $L = \sup\{x\colon x \text{ is in the support of } \rho\}$. We assume that there exists a symmetric, nonconstant, convex function h on $[-L, L]$ such that

$$g(x) \leqslant h(x) \qquad \text{for} \quad x \in [-L, L] \tag{1.1}$$

$$\int_{[-L,L]} \exp[\alpha h(x)]\, \rho(dx) < \infty \qquad \text{for all} \quad \alpha > 0 \tag{1.2}$$

The hypotheses are satisfied if, for example, g is an even polynomial with positive coefficients and ρ has bounded support. This choice of g corresponds to k-body interactions, $k \in \{2, 4,..., 2d\}$, where $2d$ is the degree of g. The hypotheses are also satisfied if, for example, g satisfies (a) and ρ has bounded support.

We note that with the definition of L in Hypothesis 1.1(c) the interval $[-L, L]$ is the smallest closed interval containing the support of ρ.

All of the results in this paper can be generalized to the case where g is C^2 and satisfies a certain two-sided real analyticity condition. For such a g, we have an example of a generalized Curie–Weiss model with infinitely many phase transitions. See Example 5.3 for details.

The *generalized Curie–Weiss model* is defined by the sequence of probability measures on $\mathbb{R}^n$, $n \in \{1, 2,...\}$, given by

$$P_{n,\beta}\{dx_1,..., dx_n\} = \exp\left[n\beta g\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)\right] \prod_{i=1}^{n} \rho(dx_i) \frac{1}{Z_n(\beta)} \tag{1.3}$$

In this formula, n is a positive integer, β is a positive real number representing the inverse absolute temperature, and $Z_n(\beta)$ is the normalization

$$Z_n(\beta) = \int_{\mathbb{R}^n} \exp\left[n\beta g\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)\right] \prod_{i=1}^{n} \rho(dx_i) \tag{1.4}$$

According to Hypothesis 1.1(a), g is nonnegative and thus, according to Hypothesis 1.1(c), h is nonnegative. Since h is convex,

$$1 \leqslant Z(n, \beta) \leqslant \left[\int_{[-L, L]} \exp[\beta h(x)] \, \rho(dx) \right]^n < \infty \tag{1.5}$$

Hence $Z(n, \beta)$ is finite.

The classical Curie–Weiss model is defined by (1.3) with $g(x) = \mathscr{J}_0 x^2/2$ and $\rho(dx) = \frac{1}{2}(\delta_1 + \delta_{-1})$. Here $\mathscr{J}_0$ is a positive constant representing the interaction strength.

In formula (1.3), we think of ρ as the distribution of a single scalar magnetic spin in the absence of interactions and of g as the interaction potential of the mean spin value. Formula (1.3) gives the joint distribution of n interacting spins at the sites $1, 2, \ldots, n$ of $\mathbb{Z}$. In the sequel, we will refer to g as the *interaction function* and to ρ as the *single-site distribution.*

The generalized Curie–Weiss model with a quartic interaction function was used by Mouritsen *et al.*[15] in order to study Ising-type models with four-body spin interactions. Limit theorems for the generalized Curie–Weiss model with a special class of interaction functions have been studied by Jeon.[12,13] These limit theorems generalize some results of Ellis and Newman.[6,7] Thermodynamic quantities calculated in the classical Curie–Weiss model coincide with those calculated via the phenomenological theory of ferromagnetism called mean field theory. For an overview of results valid for the classical Curie–Weiss model, see Ellis[4] or Thompson.[23] The classical Curie–Weiss model is also known as the Husimi–Temperley model.[11,24]

The key function used in analyzing the generalized Curie–Weiss model is the specific Gibbs free energy $\psi(\beta)$, which is defined by the formula

$$-\beta\psi(\beta) = \lim_{n \to \infty} \frac{1}{n} \log Z_n(\beta), \qquad \beta > 0 \tag{1.6}$$

For the classical Curie–Weiss model, the well-known Gaussian transform trick of Kac[14] yields

$$-\beta\psi(\beta) = \sup_{t \in \mathbb{R}} \left\{ \log \cosh t - \frac{t^2}{2\beta \mathscr{J}_0} \right\} \tag{1.7}$$

However, for the generalized Curie–Weiss model with a nonquadratic interaction function this trick is not available. Instead, we appeal to the theory of large deviations for the evaluation of the limit in (1.6).

Let S_n be the nth partial sum of i.i.d. random variables distributed by ρ. The specific Gibbs free energy $\psi(\beta)$ can be written in the form

$$
\begin{aligned}
-\beta\psi(\beta) &= \lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta) \\
&= \lim_{n\to\infty} \frac{1}{n} \log \int_{\mathbb{R}^n} \exp\left[n\beta g\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)\right] \prod_{i=1}^{n} \rho(dx_i) \\
&= \lim_{n\to\infty} \frac{1}{n} \log \textstyle\int_{\mathbb{R}} \exp[n\beta g(u)]\, P\left\{\frac{S_n}{n} \in du\right\} \qquad (1.8)
\end{aligned}
$$

We define the function

$$c(t) = \log \int_{\mathbb{R}} \exp(tx)\, \rho(dx), \qquad t \in \mathbb{R} \qquad (1.9)$$

which is the cumulant generating function of ρ. The entropy function of ρ is defined to be the Legendre–Fenchel transform $i(u)$ of $c(t)$:

$$i(u) = c^*(u) = \sup_{t\in\mathbb{R}} \{tu - c(t)\} \qquad (1.10)$$

The function $i(u)$ is strictly positive for all u not equal to the mean 0 of the symmetric probability measure ρ. According to the theory of large deviations, if A is any Borel subset of $\mathbb{R}$ whose closure does not contain the mean 0 of ρ, then the probability $P\{S_n/n \in A\}$ converges to 0 exponentially fast like $\exp[-ni(A)]$, where

$$i(A) = \inf_{u\in A} i(u) > 0 \qquad (1.11)$$

We summarize this fact by the heuristic formula

$$P\left\{\frac{S_n}{n} \in du\right\} \approx \exp[-ni(u)]\, du \qquad \text{as} \quad n\to\infty \qquad (1.12)$$

If this is inserted in (1.8), the latter suggests by Laplace's method that $\psi(\beta)$ can be determined by the variational formula

$$-\beta\psi(\beta) = \sup_{u\in\mathbb{R}} \{\beta g(u) - i(u)\} \qquad (1.13)$$

As we discuss in Section 2, formula (1.13) is rigorously true. In the case of the classical Curie–Weiss model, the right-hand side of (1.13) can be directly shown to be equal to the right-hand side of (1.7), using properties of Legendre–Fenchel transforms.

It is well known that for the classical Curie–Weiss model the supremum in (1.7) is attained at a unique nonnegative point t^*; equivalently, the supremum in (1.13) is attained at a unique nonnegative point u^* and $u^* = \tanh t^*$. This point coincides with the value of the spontaneous magnetization $m(\beta)$ for the model at that value of β. For $0 < \beta \leqslant 1$, $m(\beta) = 0$ while for $\beta > 1$, $m(\beta) > 0$. The function $m(\beta)$ is a real analytic function of $\beta \in (0, 1) \cup (1, \infty)$, but cannot be represented as the restriction of one real analytic function in any neighborhood of $\beta = 1$ [$\lim_{\beta \downarrow 1} m'(\beta) = \infty$]. The value $\beta = 1$ is called a critical value for the model. The following theorem considers the analogous situation for the generalized Curie–Weiss model. Theorem 3.1 below provides further information about the spontaneous magnetization.

Theorem 1.2. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then there exists a nonempty set $\mathscr{P}$ of positive points $\{\beta_i\}$, called critical values, which are either finite in number ($0 < \beta_1 < \cdots < \beta_N$, some $N \in \{1, 2,...\}$) or countably infinite ($0 < \beta_1 < \beta_2 < \cdots$) and divergent to ∞. This set of critical values has the following properties.

(a) There exists a function

$$m(\beta) = (0, \infty)\backslash\mathscr{P} \to [0, L) \tag{1.14}$$

such that $m(\beta) = 0$ for $0 < \beta < \beta_1$; $m(\beta) > 0$ and is strictly increasing for $\beta \in (\beta_1, \infty)\backslash\mathscr{P}$. The function $m(\beta)$ is real analytic on each connected subinterval of the set $(0, \infty)\backslash\mathscr{P}$, but cannot be represented as the restriction of one real analytic function in any neighborhood of a critical value β_i, $i \geqslant 1$. We call $m(\beta)$ the spontaneous magnetization.

(b) For $0 < \beta < \beta_1$, the supremum in the formula

$$-\beta\psi(\beta) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} \tag{1.15}$$

is attained at the unique point $u = 0$. For $\beta \in (\beta_1, \infty)\backslash\mathscr{P}$, the supremum in (1.15) is attained at the unique points $u = m(\beta)$ and $u = -m(\beta)$.

Part (b) of Theorem 1.2 has an interesting probabilistic application. For any bounded, continuous function f, consider the limit

$$\begin{aligned} &\lim_{n \to \infty} \int_{\mathbb{R}^n} f\left(\sum_{i=1}^{n} \frac{x_i}{n}\right) P_{n,\beta}(dx_1,..., dx_n) \\ &\quad = \lim_{n \to \infty} \frac{\int_{\mathbb{R}} f(u) \exp[n\beta g(u)]\, P\{S_n/n \in du\}}{\int_{\mathbb{R}} \exp[n\beta g(u)]\, P\{S_n/n \in du\}} \end{aligned} \tag{1.16}$$

The probability measures $P_{n,\beta}$ define the generalized Curie–Weiss model and are given in (1.3); S_n is the nth partial sum of i.i.d. random variables distributed by ρ. The heuristic large-derivation formula

$$P\{S_n/n \in du\} \approx \exp[-ni(u)]\,du \qquad \text{as} \quad n \to \infty \tag{1.17}$$

suggests that the limit in (1.16) should be determined by the points u at which the function $\beta g(u) - i(u)$ attains its supremum over $\mathbb{R}$. For $\beta \in (0, \infty)\backslash\mathscr{P}$ these points were determined in part (b) of Theorem 1.2. Because of symmetry, we are led to the following limit in (1.16):

$$\lim_{n\to\infty} \int_{\mathbb{R}^n} f\left(\sum_{i=1}^{n} \frac{x_i}{n}\right) P_{n,\beta}(dx_1,..., dx_n)$$
$$= \begin{cases} f(0) & \text{for} \quad \beta \in (0, \beta_1) \\ \frac{1}{2}f(m(\beta)) + \frac{1}{2}f(-m(\beta)) & \text{for} \quad \beta \in (\beta_1, \infty)\backslash\mathscr{P} \end{cases} \tag{1.18}$$

This limit is in fact rigorously true [see Theorem 3.1(e)]. For $\beta \in (0, \beta_1)$ we are in the region of relatively weak interactions among the spins, and according to (1.18), a law of large numbers for the spin per site is valid. For $\beta \in (\beta_1, \infty)$ we are in the region of relatively strong interactions among the spins, and according to (1.18), for $\beta \in (\beta_1, \infty)\backslash\mathscr{P}$ the law of large numbers breaks down. Concerning the evaluation of the limit when β is a critical value, see Remarks 1 and 2 following Theorem 3.1.

According to part (a) of Theorem 1.2, the spontaneous magnetization $m(\beta)$ cannot be represented as the restriction of one real analytic function in any neighborhood of a critical value β_i, $i \geqslant 1$. We make a definition in order to distinguish between the two possible types of nonanalytic behavior that $m(\beta)$ may have in the neighborhood of a critical value. The generalized Curie–Weiss model is said to have a *first-order phase transition* at a critical value β_i if

$$m^-(\beta_i) = \lim_{\beta \uparrow \beta_i} m(\beta) < m^+(\beta_i) = \lim_{\beta \downarrow \beta_i} m(\beta) \tag{1.19}$$

In this case β_i is called a *first-order critical value.* The generalized Curie–Weiss model is said to have a *second-order phase transition* at a critical value β_i if

$$m^-(\beta_i) = m^+(\beta_i) \qquad \text{and} \qquad \lim_{\beta \downarrow \beta_i} m'(\beta) = \infty \tag{1.20}$$

In this case β_i is called a *second-order critical value.*

It is well known that the classical Curie–Weiss model has a second-order critical value $\beta_1 = 1/\mathscr{J}_0$ and that this is the only critical value for the model. The generalized Curie–Weiss model can be more complicated. In fact, it can have a sequence of critical values with both first-order and second-order phase transitions occurring. In Section 4 a number of criteria are given for the occurrence of first-order or second-order phase transitions at the critical values. Here is a sample of these criteria:

(a) If $\beta_1 < \lim_{x \to 0} i'(x)/g'(x)$, then β_1 is a first-order critical value.

(b) If $\beta_1 = \inf\{i'(x)/g'(x)\colon 0 < x < L\}$, then β_1 is a second-order critical value.

(c) β_1 is a second-order critical value if and only if $m^+(\beta_1) = 0$, and then $g''(0) > 0$ and $\beta_1 = 1/[\int x^2\, d\rho\; g''(0)]$.

(d) If $g''(0) = 0$, then β_1 is a first-order critical value.

(e) The generalized Curie–Weiss model does not have first-order phase transitions if and only if

$$i''(x)/i'(x) \geqslant g''(x)/g'(x) \qquad \text{for all} \quad 0 < x < L \qquad (1.21)$$

(f) The generalized Curie–Weiss model has exactly one second-order phase transition at $\beta_1 = 1/[\int x^2\, d\rho \cdot g''(0)] < \infty$ if and only if there is strict inequality in (1.21) for all $0 < x < L$.

Criteria (a)–(d) are part of Theorem 4.4; criteria (e) and (f) are parts of Theorem 4.6 and Corollary 4.7, respectively. Criteria (a)–(f) and others are developed in Section 4 as consequences of a unified approach. We consider the results in this section to be among the main contributions of this paper.

In Section 5 three examples of generalized Curie–Weiss models are presented. The first model exhibits a single first-order phase transition. The second model exhibits three phase transitions with both first-order and second-order phase transitions appearing. The last model shows a cascade of infinitely many phase transitions. These examples are analyzed numerically and by means of the theorems in this paper.

In Sections 6 and 7 we prove Theorem 3.1, which states the existence and properties of the spontaneous magnetization in the generalized Curie–Weiss model.

We end this introduction by mentioning an interesting problem related to these models. An important application of the classical Curie–Weiss model is that it provides rigorous bounds to a number of quantities of interest in other models of ferromagnetism. Let β_c and $m(\beta)$ denote the critical inverse temperature and spontaneous magnetization, respectively, for any ferromagnetic model on $\mathbb{Z}^D$ with two-body interactions described by a summable ferromagnetic interaction J. Denote by β_c^{CW} and $m^{\mathrm{CW}}(\beta)$

the corresponding quantities for the classical Curie–Weiss model with interaction strength $\mathscr{J}_0 = \sum_{k \in \mathbb{Z}^D} J(k)$. We then have the result that

$$\beta_c \geqslant \beta_c^{\text{CW}} = 1/\mathscr{J}_0 \qquad \text{and} \qquad m^{\text{CW}}(\beta) \geqslant m(\beta) \qquad \text{for all} \quad \beta > 0 \tag{1.22}$$

These and related Curie–Weiss bounds have been derived in many papers.[1,9,10,16–20,23,24] An interesting problem for future research is to investigate whether quantities in ferromagnetic models with multibody interactions can be rigorously bounded by corresponding quantities in appropriately chosen generalized Curie–Weiss models.

2. SPECIFIC GIBBS FREE ENERGY

We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Our first task is the evaluation of the specific Gibbs free energy for the generalized Curie–Weiss model. The *specific Gibbs free energy* is the function $\psi(\beta)$ defined b the formula

$$\begin{aligned} -\beta\psi(\beta) &= \lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta) \\ &= \lim_{n\to\infty} \frac{1}{n} \log \int_{\mathbb{R}^n} \exp\left[n\beta g\left(\sum_{i=1}^n \frac{x_i}{n} \right) \right] \prod_{i=1}^n \rho(dx_i) \\ &= \lim_{n\to\infty} \frac{1}{n} \log \int_{\mathbb{R}} \exp[n\beta g(u)]\, P\left\{ \frac{S_n}{n} \in du \right\} \end{aligned} \tag{2.1}$$

where S_n is the nth partial sum of i.i.d. random variables distributed by ρ. In order to evaluate this limit, we appeal to the theory of large deviations.

We first define the function

$$c(t) = \log \int_{\mathbb{R}} \exp(tx)\, \rho(dx) = \log \int_{[-L,L]} \exp(tx)\, \rho(dx), \qquad t \in \mathbb{R} \tag{2.2}$$

Clearly, $c(t) > -\infty$. We now show that $c(t) < \infty$. According to Hypotheses 1.1(a) and (c), the function h on $[-L, L]$ is symmetric, nonconstant, convex, and nonnegative. Hence, there exist constants $\gamma_1 > 0$, $\gamma_2 \geqslant 0$ such that $h(x) \geqslant \gamma_1 |x| - \gamma_2$ for all $x \in [-L, L]$. It follows from (1.2) that for any $t \in \mathbb{R}$

$$\int_{[-L,L]} \exp(tx)\, \rho(dx) \leqslant \int_{[-L,L]} \exp\{|t| [h(x) + \gamma_2]/\gamma_1\}\, \rho(dx) \tag{2.3}$$

Thus, $c(t)$ is finite. The function $c(t)$ is a finite, strictly convex function on $\mathbb{R}$ and is real analytic on $\mathbb{R}$.

We define the *entropy function* of the measure ρ as the Legendre–Fenchel transform of $c(t)$:

$$i(u) = c^*(u) = \sup_{t \in \mathbb{R}} \{tu - c(t)\}, \qquad u \in \mathbb{R} \tag{2.4}$$

The function $i(u)$ is well-defined and is a convex function which maps $\mathbb{R}$ into $[0, \infty]$. For each positive integer n, let S_n be the nth partial sum of i.i.d. random variables distributed by ρ. The following well-known properties of $i(u)$ will be used in the sequel. For proofs, see Sections VII.5 and VIII.3 of ref. 4. Properties (a)–(d) state that the distributions $P\{S_n/n \in \cdot\}$ have a large-deviation property with entropy function $i(u)$.

2.1. Properties of $i(u)$

(a) $i(u)$ is lower semicontinuous.

(b) For each b real, the set $\{u \in \mathbb{R}: i(u) \leqslant b\}$ is compact.

(c) For each closed set K in $\mathbb{R}$

$$\limsup_{n \to \infty} \frac{1}{n} \log P\left\{\frac{S_n}{n} \in K\right\} \leqslant -\inf_{u \in K} i(u) \tag{2.5}$$

(d) For each open set G in $\mathbb{R}$

$$\liminf_{n \to \infty} \frac{1}{n} \log P\left\{\frac{S_n}{n} \in G\right\} \geqslant -\inf_{u \in G} i(u) \tag{2.6}$$

(e) Define $L = \sup\{x: x \text{ is in the support of } \rho\}$. The function $i(u)$ is real analytic on $(-L, L)$ and is even.

(f) $i(u) = \infty$ for $|u| > L$ and $i(L) = i(-L)$ is finite if and only if ρ has an atom at L $(\rho\{L\} > 0)$.

(g) $i(u)$ is strictly decreasing for $u \in (-L, 0]$ and is strictly increasing for $u \in [0, L)$; $i(0) = 0$. Thus, $i(u) > i(0) = 0$ for all $u \neq 0$.

(h) $\lim_{u \uparrow L} i'(u) = \infty$ and $\lim_{u \downarrow -L} i'(u) = -\infty$.

(i) $i'(u)$ defines a one-to-one mapping of $(-L, L)$ onto $\mathbb{R}$, and i' is the inverse function of c'.

(j) $i'(0) = 0$ and $i''(0) = 1/\int x^2 \, d\rho$.

In the next theorem, we prove the variational formula

$$-\beta\psi(\beta) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} \tag{2.7}$$

together with related facts.

Theorem 2.2. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. The following conclusions hold.

(a) The limit defining the specific Gibbs free energy exists and is finite and for each $\beta > 0$

$$-\beta\psi(\beta) = \lim_{n \to \infty} \frac{1}{n} \log Z_n(\beta) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} \tag{2.8}$$

(b) The function $\beta \to -\beta\psi(\beta)$ is a continuous convex function of $\beta \in [0, \infty)$ and

$$0 \leqslant -\beta\psi(\beta) \leqslant \log \int_{[-L, L]} \exp[\beta h(x)]\, \rho(dx) < \infty \tag{2.9}$$

where h is given in Hypothesis 1.1(c).

(c) For each $\beta > 0$,

$$\beta g(u) - i(u) \to -\infty \qquad \text{as} \quad |u| \to \infty \tag{2.10}$$

(d) For each $\beta > 0$, there exists $m \in (-L, L)$ such that

$$\beta g(m) - i(m) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta) \tag{2.11}$$

where $L = \sup\{x\colon x \text{ is in the support of } \rho\}$. If $L < \infty$, then the supremum in (2.11) is not attained at $m \in (-\infty, -L] \cup [L, \infty)$. It follows that if the supremum in (2.11) is attained at $m \in \mathbb{R}$, then $m \in (-L, L)$ and $\beta g'(m) = i'(m)$.

Proof. (a) As we noted in (2.1),

$$-\beta\psi(\beta) = \lim_{n \to \infty} \frac{1}{n} \log \int_{\mathbb{R}} \exp[n\beta g(u)]\, P\left\{\frac{S_n}{n} \in du\right\} \tag{2.12}$$

The large-deviation property of the distributions of S_n/n and Theorem II.7.1 in ref. 4 yield part (a), provided we show

$$\lim_{M \to \infty} \limsup_{n \to \infty} \frac{1}{n} \log \int_{\{u\colon \beta g(u) \geqslant M\}} \exp[n\beta g(u)]\, P\left\{\frac{S_n}{n} \in du\right\} = -\infty \tag{2.13}$$

Write $S_n = \sum_{i=1}^{n} X_i$, where $X_1, \ldots, X_n$ are i.i.d. random variables distributed by ρ. Using the bound $g(u) \leqslant h(u)$ for $u \in [-L, L]$, the convexity of h, and

Chebyshev's inequality, we have the following upper bound on the integral in (2.13):

$$\begin{aligned}\int_{\{u:\,\beta g(u)\geqslant M\}} &\exp[n\beta g(u)]\, P\left\{\frac{S_n}{n}\in du\right\}\\ &\leqslant \sum_{\gamma=1}^{\infty} \exp[n(\gamma+1)M]\, P\left\{\beta g\left(\frac{S_n}{n}\right)\geqslant \gamma M\right\}\\ &\leqslant \sum_{\gamma=1}^{\infty} \exp[n(\gamma+1)M]\, P\left\{\frac{\beta}{n}\sum_{j=1}^{n} h(X_j)\geqslant \gamma M\right\}\\ &\leqslant \sum_{\gamma=1}^{\infty} \exp[-n(2\gamma-1)M]\, K(\beta)^n \end{aligned}\tag{2.14}$$

where

$$K(\beta)=\int_{[-L,L]} \exp[3\beta h(u)]\,\rho(du)<\infty$$

by (1.2). The limit (2.13) follows.

(b) The inequalities in (2.9) follow immediately from (1.5). The function $\beta\to -\beta\psi(\beta)$ is a convex function of $\beta>0$ by Hölder's inequality, and it is therefore a continuous function of $\beta>0$. The continuity for $\beta\to 0^+$ follows from (2.9).

(c) We prove (2.10), first under the assumption that g is bounded at $\pm\infty$. The entropy function $i(u)$ is nonnegative, convex, and not identically zero. Hence $i(u)\to\infty$ as $|u|\to\infty$, and $\beta g(u)-i(u)\to -\infty$ as $|u|\to\infty$. We now assume that g is unbounded at ∞ and that

$$\limsup_{|u|\to\infty}[\beta g(u)-i(u)]\geqslant A \tag{2.15}$$

for some real number A. It follows that

$$-(\beta+1)\,\psi(\beta+1)=\sup_{u\in\mathbb{R}}\{g(u)+\beta g(u)-i(u)\}=\infty \tag{2.16}$$

This contradiction to part (a) of the theorem proves (2.10).

(d) It follows from (2.11) that for each $\beta>0$ the upper semicontinuous function $u\to\beta g(u)-i(u)$ attains its supremum at some point $m\in\mathbb{R}$. For $L<\infty$, since $i(u)=\infty$ for $|u|>L$, the supremum is not attained at $m\in(-\infty,-L)\cup(L,\infty)$. Suppose that $L<\infty$ and

$$\beta g(L)-i(L)=\sup_{u\in\mathbb{R}}\{\beta g(u)-i(u)\} \tag{2.17}$$

Since $i(u)$ is convex, $i'(u) \to \infty$ as $u \to L^-$, and $g'(L)$ exists, we have for all sufficiently small $\varepsilon > 0$

$$\frac{i(L) - i(L-\varepsilon)}{\varepsilon} \geqslant i'(L-\varepsilon) > \beta \frac{g(L) - g(L-\varepsilon)}{\varepsilon} \tag{2.18}$$

and so

$$\beta g(L-\varepsilon) - i(L-\varepsilon) > \beta g(L) - i(L) \tag{2.19}$$

This contradicts (2.17). It follows that if the function $u \to \beta g(u) - i(u)$ attains its supremum at $m \in \mathbb{R}$, then $m \in (-L, L)$; by calculus, $\beta g'(m) = i'(m)$. This completes the proof of the theorem. ∎

In the next section, we turn to one of the main points of this paper, which is the structure of the set of points m at which the supremum in (2.11) is attained.

3. EXISTENCE OF CRITICAL VALUES AND PROPERTIES OF THE SPONTANEOUS MAGNETIZATION

The bulk of this paper is concerned with the structure of the set of points m satisfying

$$\beta g(m) - i(m) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta) \tag{3.1}$$

We introduce the set

$$\mathscr{M} = \{(\beta, m) \in (0, \infty) \times \mathbb{R}: \beta g(m) - i(m) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta)\} \tag{3.2}$$

and its cuts at $\beta > 0$

$$\mathscr{M}_\beta = \{m \in \mathbb{R}: (\beta, m) \in \mathscr{M}\} \tag{3.3}$$

According to part (d) of Theorem 2.2, $\mathscr{M}_\beta$ is nonempty for each $\beta > 0$. Before discussing the structure of these sets for the generalized Curie–Weiss model, we review the situation for the classical Curie–Weiss model. To ease the notation, we set $\mathscr{J}_0 = 1$.

The entropy function corresponding to $\rho = \frac{1}{2}(\delta_1 + \delta_{-1})$ is given by

$$i(u) = \begin{cases} \dfrac{1-u}{2}\log(1-u) + \dfrac{1+u}{2}\log(1+u) & \text{for} \quad |u| \leqslant 1 \\ \infty & \text{for} \quad |u| > 1 \end{cases} \tag{3.4}$$

where $0 \log 0 = 0$. For the classical Curie–Weiss model with $\mathscr{J}_0 = 1$,

$$-\beta\psi(\beta) = \sup_{u \in \mathbb{R}} \{\beta u^2/2 - i(u)\} \tag{3.5}$$

The supremum is attained at points m satisfying

$$(\beta m^2/2 - i(m))' = 0 \qquad \text{or} \qquad \beta m = i'(m) = \tanh^{-1} m \tag{3.6}$$

For $0 < \beta \leqslant 1$, (3.6) has a unique solution $m = 0$, and for $\beta > 1$ three solutions $m = m(\beta)$, $-m(\beta)$, 0, where $m(\beta)$ is positive and is strictly increasing to 1 as $\beta \to \infty$. For $0 < \beta \leqslant 1$, the supremum in (3.1) is attained at $m = 0$. For $\beta > 1$, the supremum in (3.1) is attained at $m = m(\beta)$, $-m(\beta)$, not at $m = 0$. Thus, the sets $\mathscr{M}_\beta$ are given by

$$\mathscr{M}_\beta = \begin{cases} \{0\} & \text{for } \ 0 < \beta \leqslant 1 \\ \{m(\beta), -m(\beta)\} & \text{for } \ \beta > 1 \end{cases} \tag{3.7}$$

We extend the definition of $m(\beta)$ by setting $m(\beta) = 0$ for $0 < \beta \leqslant 1$. On the intervals $(0, \beta_1)$ and (β_1, ∞) the function $m(\beta)$ is real analytic. The function $m(\beta)$ is continuous at $\beta = 1$ $[\lim_{\beta \to 1} m(\beta) = 0]$, but is not differentiable at $\beta = 1$. In fact,

$$m(\beta) \sim [3(\beta - 1)]^{1/2} \qquad \text{as} \quad \beta \to 1^+ \tag{3.8}$$

The classical Curie–Weiss model with $\mathscr{J}_0 = 1$ is said to have a critical value at $\beta = 1$. The quantity $m(\beta)$, positive for $\beta > 1$, is called the spontaneous magnetization.

The next theorem considers the generalized Curie–Weiss model. The analogue of (3.7)—i.e., the structure of the sets $\mathscr{M}_\beta$ for $\beta > 0$—holds for all but at most countably many values of $\beta > 0$. The following theorem is proved in Section 7.

Theorem 3.1. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then there exists a nonempty set $\mathscr{P}$ of positive points $\{\beta_i\}$, called critical values, which are either finite in number ($0 < \beta_1 < \cdots < \beta_N$, some $N \in \{1, 2,...\}$) or countably infinite ($0 < \beta_1 < \beta_2 < \cdots$) and divergent to ∞. This set of critical values has the following properties.

(a) There exists a function

$$m(\beta)\colon (0, \infty)\backslash\mathscr{P} \to [0, L) \tag{3.9}$$

such that $m(\beta) = 0$ for $\beta \in (0, \beta_1)$; $m(\beta) > 0$ and is strictly increasing for $\beta \in (\beta_1, \infty)\backslash\mathscr{P}$. The function $m(\beta)$ is real analytic on each connected subinterval of the set $(0, \infty)\backslash\mathscr{P}$, but cannot be represented as the restriction of one real analytic function in any neighborhood of a critical value β_i, $i \geqslant 1$. We call $m(\beta)$ the spontaneous magnetization.

(b) We have

$$\mathscr{M}_\beta = \begin{cases} \{0\} & \text{for } \beta \in (0, \beta_1) \\ \{m(\beta), -m(\beta)\} & \text{for } \beta \in (\beta_1, \infty)\backslash\mathscr{P} \end{cases}$$

(c) The smallest critical value β_1 is characterized by the formulas

$$\beta_1 = \sup\{\beta > 0: \mathscr{M}_\beta = \{0\}\} = \inf\{\beta > 0: 0 \notin \mathscr{M}_\beta\} \tag{3.10}$$

(d) We have

$$\beta g'(m(\beta)) = i'(m(\beta)) \qquad \text{for} \quad \beta \in (0, \infty)\backslash\mathscr{P} \tag{3.11}$$

and

$$m(\beta) \uparrow L \qquad \text{as} \quad \beta \to \infty \qquad \text{in the set} \quad (0, \infty)\backslash\mathscr{P} \tag{3.12}$$

(e) For any bounded continuous function f

$$\lim_{n \to \infty} \int_{\mathbb{R}^n} f\left(\sum_{i=1}^n \frac{x_i}{n}\right) P_{n,\beta}(dx_1,..., dx_n)$$
$$= \begin{cases} f(0) & \text{for } \beta \in (0, \beta_1) \\ \frac{1}{2}f(m(\beta)) + \frac{1}{2}f(-m(\beta)) & \text{for } \beta \in (\beta_1, \infty)\backslash\mathscr{P} \end{cases} \tag{3.13}$$

(f) If $L < \infty$ and $\rho\{L\} > 0$, then there are only finitely many critical values $0 < \beta_1 < \cdots < \beta_N$, for some $N \in \{1, 2,...\}$.

In the next section, we discuss the possible types of nonanalytic behavior that the spontaneous magnetization $m(\beta)$ may have in the neighborhood of a critical value. We now end this section with several remarks.

Remarks. 1. Part (e) of Theorem 3.1 states a limit for the spin per site that for $\beta \in (\beta_1, \infty)\backslash\mathscr{P}$ represents a breakdown in the law of large numbers. Analogous limits also hold in the case where, for β a critical value, $\mathscr{M}_\beta$ consists of more than one pair of symmetric points. These limits can be derived by the methods of refs. 6 and 8. The methods of these papers can also be used to study the fluctuations of the spin per site (central limit theorem and related results) in the generalized Curie–Weiss model.

2. We recall from the Introduction the definition of a first-order critical value and a second-order critical value. By the methods of this

paper, one may show the following additional facts (proofs omitted). If β_i is a first-order critical value, then $\mathscr{M}_{\beta_i}$ may consist of more than two points, but $\mathscr{M}_{\beta_i}$ is always a symmetric subset of

$$[-m^+(\beta_i), -m^-(\beta_i)] \cup [m^-(\beta_i), m^+(\beta_i)] \tag{3.14}$$

which contains the endpoints

$$m^-(\beta_i) = \lim_{\beta \uparrow \beta_i} m(\beta) < \lim_{\beta \downarrow \beta_i} m(\beta) = m^+(\beta_i)$$

If β_i is a second-order critical value, then

$$\mathscr{M}_{\beta_i} = \{m^+(\beta_i), -m^+(\beta_i)\} \tag{3.15}$$

and for $\beta = \beta_i$ the limit (3.13) holds with $m(\beta)$ replaced by $m^+(\beta_i)$.

3. The specific Gibbs free energy $\psi(\beta)$ is real analytic on each connected subinterval of the set $(0, \infty)\backslash\mathscr{P}$. If β_i is a first-order critical value, then $\psi(\beta)$ is continuous, but not $\mathscr{C}^1$ in a neighborhood of β_i [$\psi'(\beta)$ has a jump discontinuity at β_i]. If β_i is a second-order critical value, then $\psi(\beta)$ is $\mathscr{C}^1$, but not $\mathscr{C}^2$ in a neighborhood of β_i. For any $\beta > 0$ that is not a first-order critical value, the derivative $[-\beta\psi(\beta)]'$ equals the specific energy:

$$[-\beta\psi(\beta)]' = g(m(\beta)) = \lim_{n\to\infty} \frac{1}{n} \log \int_{\mathbb{R}^n} g\left(\sum_{i=1}^n \frac{x_i}{n}\right) \rho_{n,\beta}(dx_1,..., dx_n) \tag{3.16}$$

4. The generalized Curie–Weiss does not exhibit a kth-order phase transition for $k \geqslant 3$, where the specific Gibbs free energy is C^{k-1}, but not C^k in a neighborhood of a critical value.

In the next section, we look more closely at the classification of critical values according to first or second order.

4. CRITERIA FOR FIRST-ORDER AND SECOND-ORDER PHASE TRANSITIONS

According to part (a) of Theorem 3.1, the spontaneous magnetization $m(\beta)$ cannot be represented as the restriction of one real analytic function in any neighborhood of a critical value β_i, $i \geqslant 1$. In fact, $m(\beta)$ may have one of two possible types of nonanalytic behavior at a critical value. In order to distinguish the two cases, we have the following definition.

Definition 4.1. The generalized Curie–Weiss model has a *first-order phase transition* at a critical value β_i if

$$m^-(\beta_i) = \lim_{\beta \uparrow \beta_i} m(\beta_i) < m^+(\beta_i) = \lim_{\beta \downarrow \beta_i} m(\beta_i) \tag{4.1}$$

In this case β_i is called a *first-order critical value.* The generalized Curie–Weiss model has a *second-order phase transition* at a critical value β_i if

$$m^-(\beta_i) = m^+(\beta_i) \qquad \text{and} \qquad \lim_{\beta \downarrow \beta_i} m'(\beta) = \infty \tag{4.2}$$

In this case β_i is called a *second-order critical value.*

Remarks. 1. In Definition 4.1, since $m(\beta) = 0$ for $\beta \in (0, \beta_1)$, we set $m^-(\beta_1) = 0$.

2. According to our discussion in Section 3, the classical Curie–Weiss model has a single critical value at $\beta_1 = 1$ and this critical value is of second order.

3. The definitions of first-order and second-order phase transitions given in Definition 4.1 are standard. See, for example, ref. 2, where the definitions are formulated in terms of the specific energy.

According to part (d) of Theorem 3.1, the specific magnetization $m(\beta)$ on the set $(\beta_1, \infty)\backslash\mathscr{P}$ is locally the inverse function of the real analytic function

$$b(x) = \frac{i'(x)}{g'(x)}, \qquad x \in (-L, 0) \cup (0, L) \tag{4.3}$$

By Hypothesis 1.1(a), $g'(x) \neq 0$ for $x \neq 0$. In this section we relate the kind of phase transition that occurs at a critical value—i.e., first or second order—to properties of $b(x)$. Since $b(x)$ is smooth and symmetric, it follows that no other kinds of nonanalytic behavior besides (4.1) and (4.2) can occur.

Since both g and i are even functions, b is also an even function. Since by Hypothesis 1.1(a) $g'(x) > 0 = g'(0)$ for $x > 0$, we see that $g''(0) \geqslant 0$. The function b can be extended to a real analytic function on $(-L, L)$ if and only if $g''(0) > 0$. Otherwise, b has a pole at 0. Defining $\mu_2 = \int x^2 \, d\rho$, we can thus set

$$b(0) = \begin{cases} i''(0)/g''(0) = 1/[g''(0)\,\mu_2] & \text{if} \quad g''(0) > 0 \\ +\infty & \text{if} \quad g''(0) = 0 \end{cases} \tag{4.4}$$

For future reference, we note that if $g''(0) > 0$, then

$$b'(0) = 0 \tag{4.5}$$

$$\begin{aligned} b''(0) &= \frac{i^{(\text{iv})}(0)\, g''(0) - i''(0)\, g^{(\text{iv})}(0)}{3[g''(0)]^2} \\ &= \frac{1}{3\mu_2^4\, g''(0)} \left[3\mu_2^2 - \mu_4 - \mu_2^3 \frac{g^{(\text{iv})}(0)}{g''(0)} \right] \end{aligned} \tag{4.6}$$

where $\mu_4 = \int x^4 \, d\rho$.

Before proceeding, let us first see how $b(x)$ is naturally related to the occurrence of phase transitions. A critical value β_i is distinguished by the fact that in a neighborhood of β_i the spontaneous magnetization $m(\beta)$ has nonanalytic behavior. According to Theorem 3.1(c) and Remark 2 following Theorem 3.1, for $\beta > \beta_1$ not a first-order critical value, $m(\beta)$ is the unique positive point at which the function $u \to \beta g(u) - i(u)$ attains its supremum. In fact, for such β

$$\beta g(m(\beta)) - i(m(\beta)) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta) \tag{4.7}$$

We now express the function $\beta g(u) - i(u)$, $u \geqslant 0$, in terms of $b(x)$ by writing

$$\beta g(u) - i(u) = \int_0^u [\beta - b(x)]\, v(dx) \tag{4.8}$$

where $v(dx) = g'(x)\, dx$ is a nonnegative measure on $[0, \infty)$.

If b has a pole at 0, then the integral is defined as

$$\lim_{\varepsilon \to 0} \int_\varepsilon^u [\beta - b(x)]\, v(dx) \tag{4.9}$$

which is well-defined since $b(x)\, v(dx) = i'(x)\, dx$ for $x > 0$. It follows that for fixed $\beta > 0$ the function $\beta g(u) - i(u)$, $u \geqslant 0$, attains its supremum at $u = u^*$ if and only if

$$\int_0^{u^*} [\beta - b(x)]\, v(dx) \geqslant \int_0^u [\beta - b(x)]\, v(dx) \qquad \text{for all} \quad u \geqslant 0 \tag{4.10}$$

Since in particular (4.10) applies to $\beta > \beta_1$ not a first-order critical value and $u^* = m(\beta)$, we obtain a geometric criterion for $m(\beta)$. Stating the criterion for $m(\beta)$ somewhat loosely, we can say that for $\beta > \beta_1$ not a first-order critical value, the "area" under the curve $\beta - b(x)$ for $0 \leqslant x \leqslant u$, $u \geqslant 0$, attains its supremum at the unique point $u = m(\beta)$, where "area" is taken with respect to the measure $v(x) = g'(x)\, dx$. A number of results that follow will be much more transparent if this criterion is kept in mind. We shall refer to it as the "area criterion for $m(\beta)$."

In Fig. 1, $m(\beta)$ equals m_1 if $\int_{m_1}^{m_2} [\beta - b(x)]\, v(dx) < 0$ and equals m_2 if the integral is positive. If the integral equals zero, then β is a first-order critical value and $m^-(\beta) = m_1$ and $m^+(\beta) = m_2$.

The first theorem in this section, Theorem 4.4, gives some conditions for a first-order or a second-order phase transition at the smallest critical value β_1. Theorem 4.5 describes models without first-order phase transitions. We start with two lemmas.

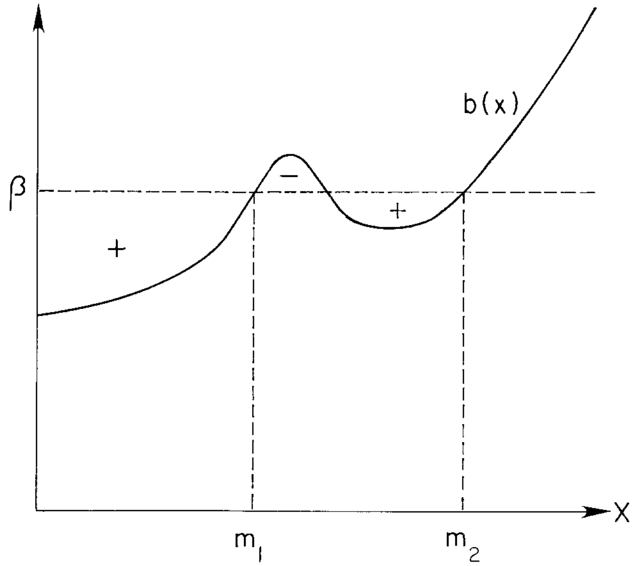

Fig. 1. Area criterion for $m(\beta)$.

Lemma 4.2. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then $\limsup_{x \uparrow L} b(x) = \infty$; in particular, $b(x)$ is nonconstant on the interval $0 < x < L$.

Proof. If L is finite, then $i'(x) \to \infty$ as $x \uparrow L$; since g' is bounded on $[-L, L]$, it follows that $b(x) \to \infty$ as $x \uparrow L$. On the other hand, if $L = \infty$, then by Theorem 2.2(c) for any $\beta > 0$ and all sufficiently large x

$$\infty > i(x)/g(x) \geqslant \beta \tag{4.11}$$

Now suppose that $\limsup_{x \to \infty} b(x) < \infty$. Then there exists a positive number A such that

$$b(x) = i'(x)/g'(x) \leqslant A \qquad \text{for all} \quad x \geqslant 1 \tag{4.12}$$

It follows from (4.11) and (4.12) that for any $\beta > 0$ and all sufficiently large x

$$\beta g(x) - i(1) \leqslant i(x) - i(1) \leqslant A[g(x) - g(1)]$$

This is impossible, since $\lim_{x \to \infty} g(x) > 0$, and so $\limsup_{x \to \infty} b(x) = \infty$. ∎

Lemma 4.3. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Define the quantity

$$\underline{\beta} = \inf\{b(x): 0 < x < L\} \tag{4.13}$$

Then we have the inequalities

$$0 \leqslant \underline{\beta} \leqslant \beta_1 \leqslant b(0) \qquad \text{and} \qquad 0 < \beta_1 \tag{4.14}$$

Proof. The quantity $\underline{\beta}$ is nonnegative, since $b(x)$ is a positive function.

We next prove that $\underline{\beta} \leqslant \beta_1$. According to part (d) of Theorem 2.2, the set $\mathscr{M}_\beta$ is nonempty for each $\beta > 0$, and if $m \in \mathscr{M}_\beta$, then $\beta g'(m) = i'(m)$. Hence, if $\mathscr{M}_\beta$ contains a nonzero m, then $b(m) = \beta$ and so $\beta \geqslant \underline{\beta}$; i.e., if $\beta < \underline{\beta}$, then $\mathscr{M}_\beta = \{0\}$. Since $\beta_1 = \sup\{\beta > 0\colon \mathscr{M}_\beta = \{0\}\}$ [Theorem 3.1(c)], we conclude that $\underline{\beta} \leqslant \beta_1$.

To show $\beta_1 \leqslant b(0)$, we may assume by (4.4) that $g''(0) > 0$. For any $\beta > b(0) = \lim_{x \to 0} i'(x)/g'(x)$, there exists $\varepsilon > 0$ such that

$$\beta g'(x) - i'(x) > 0 \qquad \text{for all} \quad 0 < |x| \leqslant \varepsilon \tag{4.15}$$

Since $g(0) = i(0) = 0$, it follows that

$$\sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} \geqslant \beta g(\varepsilon) - i(\varepsilon) > 0 = \beta g(0) - i(0)$$

Hence $0 \notin \mathscr{M}_\beta$. Since $\beta_1 = \inf\{\beta > 0 \notin \mathscr{M}_\beta\}$ [Theorem 3.1(c)], we conclude that $\beta_1 \leqslant b(0)$.

We now prove that $\beta_1 > 0$. Let us assume that $\beta_1 = 0$ and obtain a contradiction. Since $\beta_1 = \inf\{\beta > 0\colon 0 \notin \mathscr{M}_\beta\}$ and $\mathscr{M}_\beta$ is nonempty for each $\beta > 0$ [Theorem 2.2(d)], there exist sequences $1 \geqslant \gamma_n \downarrow 0$ and $m_n \in (0, L)$ such that

$$\begin{aligned} g(m_n) - i(m_n) &\geqslant \gamma_n g(m_n) - i(m_n) \\ &= \sup_{u \in \mathbb{R}} \{\gamma_n g(u) - i(u)\} \geqslant 0 \end{aligned} \tag{4.16}$$

But $b(m_n) = \gamma_n \downarrow 0$ and $\inf_{0 \leqslant x \leqslant M} b(x) > 0$ for any $M \in (0, L)$. It follows that $m_n \uparrow L$. If $L = \infty$, then part (c) of Theorem 2.2 implies that $g(m_n) - i(m_n) \to -\infty$. This contradiction to (4.16) shows that $\beta_1 > 0$. Assume now that $L < \infty$. As $m_n \uparrow L$, the quantities $\{g'(m_n)\}$ stay bounded, but $i'(m_n) \to \infty$. Hence

$$\gamma_n g'(m_n) - i'(m_n) \to -\infty$$

But this contradicts the fact that $\gamma_n g'(m_n) - i'(m_n) = 0$. Again we conclude that $\beta_1 > 0$. The proof of Lemma 4.3 is complete. ▌

According to parts (a) and (b) of the next theorem, whether the smallest critical value β_1 is a first-order critical value or a second-order critical value may be characterized by whether we have strict inequalities in (4.14) or not.

Theorem 4.4. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Define $\mu_2 = \int x^2 \, d\rho$ and $\mu_4 = \int x^4 \, d\rho$.

(a) If $\beta_1 < b(0)$, then $\underline{\beta} < \beta_1 < b(0)$ and β_1 is a first-order critical value.

(b) If $\underline{\beta} = \beta_1$, then $\underline{\beta} = \beta_1 = b(0)$ and β_1 is a second-order critical value.

(c) β_1 is a second-order critical value if and only if $m^+(\beta_1) = 0$, and then

$$\beta_1 = b(0) = \frac{1}{\mu_2 \, g''(0)} < \infty \tag{4.17}$$

(d) In particular, assume either that $g''(0) = 0$ or that $g''(0) > 0$ and

$$3\mu_2^2 < \mu_4 + \mu_2^3 \, g^{(\mathrm{iv})}(0)/g''(0) \tag{4.18}$$

[i.e., $b''(0) < 0$; see (4.6)]. Then $\beta_1 < b(0)$. Hence, by part (a), β_1 is a first-order critical value.

Remarks. 1. According to part (d) of the theorem, if g is an even polynomial with no quadratic term, then β_1 is a first-order critical value. The case $g(x) = x^4$ was considered by Mouritsen *et al.*[15]

2. Only in the case where $\underline{\beta} < \beta_1 = b(0)$ can we not *a priori* say what kind of phase transition occurs at β_1. Typically in this case a second-order phase transition at β_1 is "closely" followed by a first-order phase transition. If both happen to fall together, then the second-order phase transition is suppressed by the first-order phase transition. These remarks as well as parts (a) and (b) of Theorem 4.4 may be clarified by referring to the "area criterion for $m(\beta)$" presented earlier in this section.

Proof of Theorem 4.4. (a) We asume that $\beta_1 < b(0)$. According to Theorem 3.1, the spontaneous magnetization $m(x)$ is positive for $x \in (\beta_1, \beta_2)$ and for $x \in (\beta_1, \beta_2)$ satisfies $b(m(x)) = x$. Taking $x \downarrow \beta_1$, we see that the quantity $m^+(\beta_1) = \lim_{x \downarrow \beta_1} m(x)$ satisfies $b(m^+(\beta_1)) = \beta_1$. In general, $m^+(\beta_1) \geqslant 0$. If $m^+(\beta_1) = 0$, then

$$\beta_1 = b(m^+(\beta_1)) = b(0) \tag{4.19}$$

But this contradicts the hypothesis that $\beta_1 < b(0)$. It follows that $m^+(\beta_1) > 0$. Since $m^-(\beta_1) = 0$, β_1 must be a first-order critical value.

We now prove that $\underline{\beta} < \beta_1$. According to Theorem 3.1(b), for $\beta_1 < \beta < \beta_2$ the set $\mathscr{M}_\beta$ equals $\{m(\beta), -m(\beta)\}$, and so

$$\sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = \beta g(m(\beta)) - i(m(\beta)) > 0 = \beta g(0) - i(0) \quad (4.20)$$

Taking $\beta \downarrow \beta_1$, we see that $\beta_1 g(m^+(\beta_1)) - i(m^+(\beta_1)) \geqslant 0$ or

$$\int_0^{m^+(\beta_1)} [\beta_1 g'(x) - i'(x)] \, dx \geqslant 0 \quad (4.21)$$

Since $\beta_1 < b(0)$, for a suitable choice of $\varepsilon > 0$, $\beta_1 < b(x)$ for all $-\varepsilon < x < \varepsilon$. This implies that

$$\beta_1 g'(x) - i'(x) < 0 \qquad \text{for} \quad 0 \leqslant x < \varepsilon \quad (4.22)$$

Hence by (4.21) there exists $x_0 \in (\varepsilon, m^+(\beta_1))$ such that $\beta_1 g'(x_0) - i'(x_0) > 0$. We conclude that

$$\underline{\beta} \leqslant b(x_0) < \beta_1. \quad (4.23)$$

This completes the proof of part (a).

(b) We assume that $\underline{\beta} = \beta_1$. Then $\beta_1 \leqslant b(x)$ for all $0 < x < L$. Since $b(x)$ is a real analytic function on the interval $0 < x < L$ and is nonconstant on that interval, equality in the equation $b(x) = \beta_1$ holds at most at countably many points on the interval $0 < x < L$. Hence

$$\beta_1 g'(x) - i'(x) \leqslant 0 \qquad \text{for all} \quad 0 < x < L \quad (4.24)$$

with equality at most at countably many points. Thus, for all $0 < u < L$

$$\int_0^u [\beta_1 g'(x) - i'(x)] \, dx < 0 \quad (4.25)$$

In the proof of part (a), we showed that

$$\int_0^{m^+(\beta_1)} [\beta_1 g'(x) - i'(x)] \, dx \geqslant 0 \quad (4.26)$$

The proof of this inequality did not use the hypothesis in part (a) that $\beta_1 < b(0)$. Comparing the last two displays, we see that $m^+(\beta_1)$ must equal zero. Hence β_1 must be a second-order critical value.

We now prove that $\beta_1 = b(0)$. By Lemma 4.3, $\beta_1 \leqslant b(0)$. By part (a) of the present theorem, if $\beta_1 < b(0)$, then β_1 is a first-order critical value. However, we have just proved that β_1 is a second-order critical value. It follows that $\beta_1 = b(0)$. This completes the proof of part (b).

(c) If β_1 is a second-order critical value, then $m^+(\beta_1)=m^-(\beta_1)$. Since in general $m^-(\beta_1)=0$, it follows that $m^+(\beta_1)=0$. Now suppose that $m^+(\beta_1)=\lim_{\beta\downarrow\beta_1} m(\beta)=0$. We prove that β_1 is a second-order critical value. Since $m(\beta)>0$ for $\beta\in(\beta_1,\beta_2)$ and

$$b(m(\beta))=i'(m(\beta))/g'(m(\beta))=\beta \tag{4.27}$$

(Theorem 3.1), it follows that

$$\lim_{\beta\downarrow\beta_1} b(m(\beta))=\lim_{x\to 0} b(x) \qquad \text{exists and equals } \beta_1 \tag{4.28}$$

However,

$$\lim_{x\to 0} b(x)=\lim_{x\to 0}\frac{i'(x)/x}{g'(x)/x} \tag{4.29}$$

and $\lim_{x\to 0} i'(x)/x=i''(0)=1/\mu_2$. If $g''(0)=0$, then the limit in (4.29) would be ∞. This contradicts (4.28), proving that $g''(0)>0$. In this case, b can be extended to a real analytic function on $(-L, L)$, and we have

$$b(0)=\frac{1}{\mu_2 g''(0)}<\infty \qquad \text{and} \qquad b'(m^+(\beta_1))=b'(0)=0 \tag{4.30}$$

Since for $\beta\in(\beta_1,\beta_2)$, $m(\beta)$ is differentiable, (4.27) implies that

$$m'(\beta)=\frac{1}{b'(m(\beta))} \qquad \text{and} \qquad \lim_{\beta\downarrow\beta_1} m'(\beta)=\infty \tag{4.31}$$

It follows that β_1 is a second-order critical value. We have already shown in this proof under the hypothesis $m^+(\beta_1)=0$ that $g''(0)>0$ and b can be extended to a real analytic function on $(-L, L)$. In this case, the limit in (4.29) equals

$$\lim_{x\to 0} b(x)=b(0)=\frac{1}{\mu_2 g''(0)}<\infty \tag{4.32}$$

Combining this with (4.28) completes the proof of part (c).

(d) If $g''(0)=0$, then $b(0)=\infty$ and $\beta_1<b(0)$ holds trivially. Assume $g''(0)>0$. According to (4.6),

$$b''(0)=\frac{1}{3\mu_2^4 g''(0)}\left[3\mu_2^2-\mu_4-\mu_2^3\frac{g^{(iv)}(0)}{g''(0)}\right] \tag{4.33}$$

Hence the inequality in (4.18) is equivalent to $b''(0)<0$. Since $b'(0)=0$, it follows that for a suitable $\varepsilon>0$, $b(x)<b(0)$ for all x satisfying $0<|x|\leqslant\varepsilon$.

Since $b(x) = i'(x)/g'(x)$ and $g(0) = 0 = i(0)$, we find by integrating that $b(0)\, g(\varepsilon) - i(\varepsilon) > 0$. By continuity, $\tilde{\beta} g(\varepsilon) - i(\varepsilon) > 0$ for any $\tilde{\beta} < b(0)$, which is sufficiently close to $b(0)$. Thus,

$$\sup_{u \in \mathbb{R}} \{\tilde{\beta} g(u) - i(u)\} \geqslant \tilde{\beta} g(\varepsilon) - i(\varepsilon) > 0 = \tilde{\beta} g(0) - i(0) \tag{4.34}$$

and so $0 \notin \mathcal{M}_{\tilde{\beta}}$. Since $\beta_1 = \inf\{\beta > 0 : 0 \notin \mathcal{M}_\beta\}$, we conclude that $\beta_1 \leqslant \tilde{\beta} < b(0)$. This completes the proof of Theorem 4.4. ∎

Next, we characterize generalized Curie–Weiss models without first-order phase transitions. For this purpose, we need a definition.

Definition 4.5. An inverse absolute temperature β is *simple* if there exists at most one $x \in (0, L)$ such that $b(x) = \beta$.

Theorem 4.6. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then the following statements are equivalent.

(a) The generalized Curie–Weiss model does not have first-order phase transitions.

(b) The function $b(x) = i'(x)/g'(x)$ is strictly increasing on the interval $0 < x < L$; since $b(x)$ is real analytic, this is equivalent to $b'(x) \geqslant 0$ or

$$i''(x)/i'(x) \geqslant g''(x)/g'(x) \qquad \text{for all} \quad 0 < x < L \tag{4.35}$$

(c) All $\beta > 0$ are simple.

Proof. (a) $\Rightarrow$ (b). We assume that the generalized Curie–Weiss model does not have first-order phase transitions. According to Theorem 3.1, $m(\beta)$ is a continuous, strictly increasing function of $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$, $m(\beta) = 0$ for $0 < \beta < \beta_1$, $b(m(\beta)) = \beta$ for $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$, and $m(\beta) \uparrow L$ as $\beta \to \infty$ in the set $(0, \infty) \backslash \mathscr{P}$. Since there are no first-order critical values, we may extend the definition of $m(\beta)$ to $\beta \in \{\beta_1, \beta_2, \beta_3, \ldots\}$ by continuity:

$$m(\beta_i) = \lim_{\beta \uparrow \beta_i} m(\beta) = \lim_{\beta \downarrow \beta_i} m(\beta) \tag{4.36}$$

It follows that m maps the interval (β_1, ∞) onto the interval $(0, L)$, is strictly increasing, and equals the inverse of b restricted to $(0, L)$. Hence $b(x)$ is strictly increasing on the interval $0 < x < L$.

(b) $\Rightarrow$ (c). If some $\beta > 0$ is not simple, then there exist values $x_1 < x_2$ in $(0, L)$ such that $b(x_1) = b(x_2) = \beta$. It follows that $b(x)$ is not strictly increasing on the interval $0 < x < L$.

(c) ⇒ (a). We show that if a first-order phase transition occurs at a critical value β_i, then β_i is not simple. First, suppose that $i \geqslant 2$, so that $0 < m^-(\beta_i) < m^+(\beta_i)$. For $\beta > \beta_i$ and sufficiently close to β_i, the set $\mathscr{M}_\beta$ equals $\{m(\beta), -m(\beta)\}$ [Theorem 3.1(b)]; i.e., for all such β and all $x \in (-L, L)$,

$$\beta g(m(\beta)) - i(m(\beta)) \geqslant \beta g(x) - i(x) \tag{4.37}$$

Taking $\beta \downarrow \beta_i$, we see that the function $x \to \beta_i g(x) - i(x)$ attains its supremum on $(-L, L)$ at the interior point $m^+(\beta_i)$. Thus

$$\beta_i g'(m^+(\beta_i)) - i'(m^+(\beta_i)) = 0 \qquad \text{or} \qquad b(m^+(\beta_i)) = \beta_i \tag{4.38}$$

By a similar proof, $b(m^-(\beta_i)) = \beta_i$. It follows that β_i is not simple.

We now show that if a first-order phase transition occurs at the critical value β_1, then β_1 is not simple. We have $0 = m^-(\beta_1) < m^+(\beta_1)$, and by the same proof used in the case $i \geqslant 2$, one proves $b(m^+(\beta_1)) = \beta_1$. According to Lemma 4.3 and Theorem 4.4,

$$0 \leqslant \underline{\beta} < \beta_1 \leqslant b(0) \tag{4.39}$$

If $\beta_1 < b(0)$, then the inequality $\underline{\beta} < \beta_1$ and the fact that $\limsup_{x \uparrow L} b(x) = \infty$ (Lemma 4.2) imply that there exists a positive point $x_0 \neq m^+(\beta_1)$ such that $b(x_0) = \beta_1$. Since also $b(m^+(\beta_1)) = \beta_1$, it follows that if $\beta_1 < b(0)$, then β_1 is not simple. We now prove by contradiction that if $\beta_1 = b(0)$, then β_1 is not simple. Suppose that $m = m^+(\beta_1)$ is the unique positive point satisfying $b(m) = \beta_1$. We have

$$b(x) < \beta_1 \qquad \text{for all} \quad x \in (0, m^+(\beta_1)) \tag{4.40}$$

and so

$$\beta_1 g(m^+(\beta_1)) - i(m^+(\beta_1)) = \int_0^{m^+(\beta_1)} [\beta_1 - b(x)]\, g'(x)\, dx > 0 \tag{4.41}$$

By continuity $\beta g(m^+(\beta_1)) - i(m^+(\beta_1)) > 0$ for all $\beta < \beta_1$ that are sufficiently close to β_1. It follows that

$$\sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} \geqslant \beta g(m^+(\beta_i)) - i(m^+(\beta_1)) > 0 = \beta g(0) - i(0) \tag{4.42}$$

and thus that $0 \notin \mathscr{M}_\beta$. This contradicts the fact that $\beta_1 = \sup\{\beta > 0: \mathscr{M}_\beta = \{0\}\}$ [Theorem 3.1(c)]. We conclude that $m = m^+(\beta_1)$ is *not* the unique positive point satisfying $b(m) = \beta_1$; i.e., β_1 is not simple. This completes the proof of Theorem 4.6. ∎

The next corollary characterizes generalized Curie–Weiss models that have exactly one second-order phase transition at $\beta_1 = b(0)$ and no other phase transitions.

Corollary 4.7. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then the following statements are equivalent.

(a) The generalized Curie–Weiss model has exactly one second-order phase transition at $\beta_1 = b(0)$ and no other phase transitions.

(b) The derivative of the function $b(x) = i'(x)/g'(x)$ is positive on the interval $0 < x < L$; i.e., $b'(x) > 0$ or

$$i''(x)/i'(x) > g''(x)/g'(x) \qquad \text{for all} \quad 0 < x < L \tag{4.43}$$

Proof. (a) ⇒ (b). If (a) holds, then, as in the proof of (a) ⇒ (b) in the previous theorem, we see that m maps the interval (β_1, ∞) onto $(0, L)$. Since by assumption there are no second-order phase transitions for $\beta > \beta_1$, we have $0 < m'(\beta) < \infty$ for $\beta > \beta_1$. Since $b(m(\beta)) = \beta$ for all $\beta > \beta_1$, it follows that

$$b'(m(\beta)) = \frac{1}{m'(\beta)} > 0 \qquad \text{for all} \quad \beta > \beta_1$$

This gives (b).

(b) ⇒ (a). If (b) holds, then by the previous theorem there are no first-order phase transitions. Assume that a second-order phase transition occurs at critical values $\{\beta_2, \beta_3, ...\}$. As in the proof of the previous theorem, we may extend the definition of $m(\beta)$ to $\beta \in \{\beta_2, \beta_3, ...\}$ by continuity; for all $\beta > \beta_1$ we have $b(m(\beta)) = \beta$. Since $b'(x) > 0$ for all $0 < x < L$, the inverse function theorem guarantees that m is of class $\mathscr{C}^1$ on some neighborhood of each β_i, $i \in \{2, 3, ...\}$. This implies that a second-order phase transition cannot occur at β_i. Part (a) follows. ∎

There is an important class of models for which the strict inequality (4.44) in part (b) of Corollary 4.7 can be verified to hold. Let us assume that the interaction function g is quadratic [$g(x) = cx^2$ with $c > 0$] and the single-site distribution ρ satisfies Hypotheses 1.1 together with the single-site GHS inequality:

$$c'''(t) \leqslant 0 \qquad \text{for} \quad t \geqslant 0 \tag{4.44}$$

where $c(t) = \log \int_{\mathbb{R}} \exp(tx)\, \rho(dx)$. A large class of measures ρ is known for which (4.44) is valid; this class includes $\rho = \frac{1}{2}(\delta_1 + \delta_{-1})$.[5,7] Inequality

(4.44) states that $c'(t)$ is concave on the interval $0 \leqslant t < \infty$. It is not hard to show that since ρ satisfies Hypotheses 1.1, $c'(t)$ is in fact strictly concave on the interval $0 \leqslant t < \infty$ (see ref. 3, p. 309). It follows that the inverse function $i'(x)$ is strictly convex for $0 \leqslant x < L$ or that

$$i'(x) < i''(x) \cdot x \qquad \text{for all} \quad 0 < x < L \tag{4.45}$$

This is equivalent to the inequality

$$1/x = g''(x)/g'(x) < i''(x)/i'(x) \qquad \text{for all} \quad 0 < x < L \tag{4.46}$$

which is (4.43). This gives the next corollary.

Corollary 4.8. We assume that the interaction function g is quadratic [$g(x) = cx^2$ with $c > 0$] and the single-site distribution ρ satisfies Hypotheses 1.1 together with the single-site GHS inequality (4.44). Then there exists exactly one second-order phase transition at $\beta_1 = b(0) = 1/(c\mu_2)$ and no other phase transitions.

In the next section, we study the phase transitions in three examples of generalized Curie–Weiss models.

5. THREE EXAMPLES

In this section, we give three examples of generalized Curie–Weiss models. The first example presents a generalized Curie–Weiss model with a four-body interaction which exhibits a single first-order phase transition. A model with a higher order interaction is treated in the second example. This model exhibits three phase transitions with both first-order and second-order phase transitions appearing. A model with infinitely many phase transitions is considered in the third example. For each model the smallest critical value β_1 may be determined by the formula in Theorem 3.1(c): $\beta_1 = \sup\{\beta > 0 \colon \mathscr{M}_\beta = \{0\}\}$.

Example 5.1. We consider the generalized Curie–Weiss model with interaction function

$$g(x) = \alpha_1 x^4/4! + \alpha_2 x^2/2!, \qquad \alpha_1, \alpha_2 > 0 \tag{5.1}$$

and single-site distribution $\rho = \frac{1}{2}(\delta_1 + \delta_{-1})$. Then (3.6), (4.4), and (4.6) give

$$b(x) = \tanh^{-1} x/(\alpha_1 x^3/3! + \alpha_2 x) \tag{5.2}$$

$$b(0) = 1/\alpha_2, \qquad b''(0) = (2\alpha_2 - \alpha_1)/\alpha_2^2 \tag{5.3}$$

Part (d) of Theorem 4.4 tells us that if $\alpha_1 > 2\alpha_2$, then $\beta_1 < b(0)$ and that β_1 is a first-order phase transition. Numerical calculations for $\alpha_1 = 6$ and $\alpha_2 = 2$ show that

$$\beta_1 \approx 0.491 \qquad \text{and} \qquad m^+(\beta_1) = \lim_{\beta \downarrow \beta_1} m(\beta) \approx 0.63 \tag{5.4}$$

Let us recall the definitions

$$\mathcal{M} = \{(\beta, m) \in (0, \infty) \times \mathbb{R}: \beta g(m) - i(m) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta)\} \tag{5.5}$$

$$\mathcal{M}_\beta = \{m \in \mathbb{R}: (\beta, m) \in \mathcal{M}\} \qquad \text{for} \quad \beta > 0 \tag{5.6}$$

and define the set

$$\mathcal{M}^+ = \mathcal{M} \cap ((0, \infty) \times (0, \infty)) \tag{5.7}$$

According to Theorem 2.2(d), if $m \in \mathcal{M}_\beta$ for some $\beta > 0$ and if also $m > 0$, then $b(m) = \beta$, and so $(\beta, m) = (b(m), m) \in \mathcal{M}^+$ (see Lemma 6.1 below). The portions of the curve in Fig. 2 marked with a solid line depict the set $\mathcal{M}^+ \cup \{(\beta, 0)\colon 0 < \beta \leqslant \beta_1\}$ in Example 5.1. Each point in $\mathcal{M}^+$ corresponding to some $\beta > \beta_1$ has the form $(\beta, m(\beta))$, where $m(\beta)$ is the spontaneous magnetization. The portion of the curve in Fig. 2 marked with a dot–dash line represents all points of the form $(b(m), m)$, $m \in (0, L)$, that are not in $\mathcal{M}^+$.

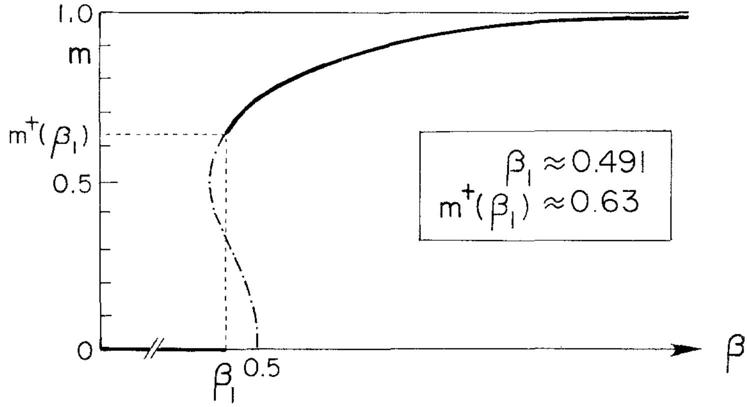

Fig. 2. This refers to Example 5.1. The portions of the curve marked with a solid line depict the set $\mathcal{M}^+ \cup \{(\beta, 0)\colon 0 < \beta \leqslant \beta_1\}$. Each point in $\mathcal{M}^+$ corresponding to some $\beta > \beta_1$ has the form $(\beta, m(\beta))$, where $m(\beta)$ is the spontaneous magnetization. The line $m = 1$ is an asymptote for $m(\beta)$ as $\beta \to \infty$. The portion of the curve marked with a dot–dash line represents all points of the form $(b(m), m)$, $m \in (0, 1)$, which are not in $\mathcal{M}^+$.

Example 5.2. We consider the generalized Curie–Weiss model with interaction function

$$g(x) = 0.15x^{60} + x^6 + 4x^2 \tag{5.8}$$

and single-site distribution $\rho = \frac{1}{2}(\delta_1 + \delta_{-1})$. Numerical calculations show that this model has three phase transitions at

$$\beta_1 = b(0) = 0.125, \qquad \beta_2 \approx 0.1313, \qquad \text{and} \qquad \beta_3 \approx 0.1445. \tag{5.9}$$

The phase transition at β_1 is second order [see Theorem 4.4(c)], while the phase transitions at β_2 and β_3 are first order. Decreasing carefully the coefficient 0.15 in the first term in (5.8), we find a number $\alpha_1 \in (0, 0.15)$ such that for the interaction function $\alpha_1 x^{60} + x^6 + 4x^2$ the third phase transition becomes one of second order at a slightly changed critical value β_3. The portions of the curve in Fig. 3 marked with a solid line depict the set $\mathscr{M}^+ \cup \{(\beta, 0)\colon\ 0 < \beta \leqslant \beta_1\}$ in Example 5.2. Each point in $\mathscr{M}^+$ corresponding to some $\beta > \beta_1$, $\beta \neq \beta_2, \beta_3$, has the form $(\beta, m(\beta))$, where $m(\beta)$ is the spontaneous magnetization. The portion of the curve in Fig. 3 marked with a dot–dash line represents all points of the form $(b(m), m)$, $m \in (0, L)$, which are not in $\mathscr{M}^+$.

Example 5.3. In this example, the interaction function g is not real analytic as required by Hypothesis 1.1(a). Instead g satisfies the following hypothesis.

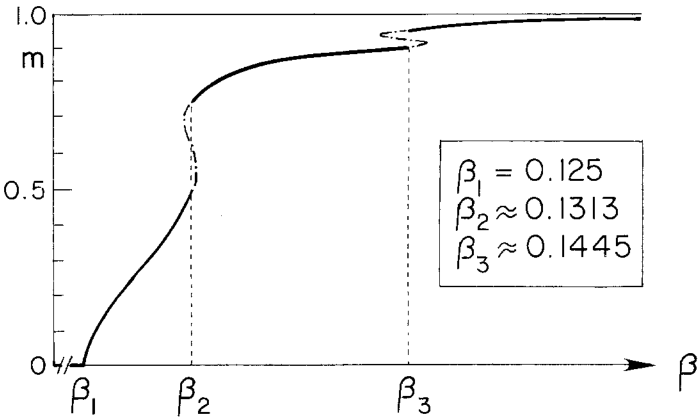

Fig. 3. This refers to Example 5.2. The portions of the curve marked with a solid line depict the set $\mathscr{M}^+ \cup \{(\beta, 0)\colon\ 0 < \beta \leqslant \beta_1\}$. Each point in $\mathscr{M}^+$ corresponding to some $\beta > \beta_1$, $\beta \neq \beta_2, \beta_3$, has the form $(\beta, m(\beta))$, where $m(\beta)$ is the spontaneous magnetization. The line $m = 1$ is an asymptote for $m(\beta)$ as $\beta \to \infty$. The portions of the curve marked with a dot–dash line represent all points of the form $(b(m), m)$, $m \in (0, 1)$, which are not in $\mathscr{M}^+$.

Hypothesis 1.1(a)′. The function g is an even, twice continuously differentiable function on $\mathbb{R}$, is strictly increasing on $[0, \infty)$ with $g(0)=0$, and is two-sided, real analytic in the sense that for each $x \in \mathbb{R}$ there exists $\delta > 0$ and two real analytic functions g_1 and g_2 on $(x-\delta, x+\delta)$ such that

$$g = g_1 \quad \text{on} \quad (x-\delta, x] \qquad \text{and} \qquad g = g_2 \quad \text{on} \quad [x, x+\delta) \tag{5.10}$$

All the results in this paper can be generalized if the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1(a)′, (b), and (c).

We now turn to our example of a generalized Curie–Weiss model with infinitely many phase transitions. We define

$$g(x) = \int_0^x g'(u)\, du \tag{5.11}$$

where

$$g'(x) = \begin{cases} (4/\pi)x & \text{for} \quad |x| \leqslant 1/2 \\ [\pi x + \sin(2\pi x)]^{-1} & \text{for} \quad |x| > 1/2 \end{cases} \tag{5.12}$$

and let ρ be the standard Gaussian distribution

$$\rho(dx) = (2\pi)^{-1/2} \exp(-x^2/2)\, dx \tag{5.13}$$

Clearly g and ρ satisfy Hypotheses 1.1(a)′, (b), and (c) [with $h(x) = \alpha\, |x|$, some $\alpha > 0$]. Since $i(x) = x^2/2$, we find

$$b(x) = \frac{i'(x)}{g'(x)} = \begin{cases} \pi/4 & \text{for} \quad |x| \leqslant 1/2 \\ \pi x^2 + x \sin 2\pi x & \text{for} \quad |x| > 1/2 \end{cases} \tag{5.14}$$

For any positive integer n,

$$b'(n+1/2) = 0 \qquad \text{and} \qquad b''(n+1/2) = -2\pi < 0 \tag{5.15}$$

while

$$b(n+1/2) = \pi(n+1/2)^2 < b(n+3/2) \tag{5.16}$$

It is now easy to see that for each positive integer n, there is a first-order critical value β_{n+1} near the value

$$b(n+1/2) = \pi(n+1/2)^2 \tag{5.17}$$

[$\beta_{n+1} < b(n+1/2)$]. Thus, the model exhibits infinitely many phase transitions.

This completes our presentation of examples. The remaining sections of the paper prove Theorem 3.1.

6. PRELIMINARY RESULTS NEEDED FOR PROOFS OF THEOREM 3.1

Let us first state some obvious consequences of Hypotheses 1.1 and recall some definitions. Since the interaction function g is even and the single-site distribution ρ is symmetric, we have

$$g'(0) = c'(0) = i'(0) = 0 \tag{6.1}$$

where $c(t) = \log \int_{\mathbb{R}} \exp(tx)\, \rho(dx)$ for t real and $i(u)$ is the Legendre–Fenchel transform of $c(t)$ [see (2.4)]. The function g is real analytic on $\mathbb{R}$ and strictly increasing on $[0, \infty)$. Thus,

$$g'(x) = -g'(-x) > 0 \qquad \text{for all} \quad x \in (0, L) \tag{6.2}$$

where $L = \sup\{x\colon x \text{ is in the support of } \rho\}$. We define

$$b(x) = i'(x)/g'(x) \qquad \text{for} \quad x \in (-L, 0) \cup (0, L) \tag{6.3}$$

The function $b(x)$ is positive, even, and real analytic on $(-L, 0) \cup (0, L)$. If $g''(0) > 0$, then b can be extended to a real analytic function on $(-L, L)$ by setting

$$b(0) = i''(0)/g''(0) = 1 \Big/ \left[g''(0) \int x^2\, d\rho \right] > 0 \tag{6.4}$$

In this case,

$$b'(0) = 0 \tag{6.5}$$

If, on the other hand, $g''(0) = 0$, then the function b has a pole at zero.

One of our tasks in Theorem 3.1 is the computation of the set

$$\mathscr{M}_\beta = \{m \in \mathbb{R}\colon (\beta, m) \in \mathscr{M}\} \qquad \text{for} \quad \beta > 0 \tag{6.6}$$

where

$$\mathscr{M} = \{(\beta, m) \in (0, \infty) \times \mathbb{R}\colon \beta g(m) - i(m) = \sup_{u \in \mathbb{R}} \{\beta g(u) - i(u)\} = -\beta\psi(\beta)\} \tag{6.7}$$

By the evenness of g and i, we may restrict our attention to the sets

$$\mathscr{M}^+ = \mathscr{M} \cap ((0, \infty) \times (0, \infty)) \tag{6.8}$$

and

$$\mathscr{M}_\beta^+ = \{m \in (0, \infty)\colon (\beta, m) \in \mathscr{M}^+\} \qquad \text{for} \quad \beta > 0 \tag{6.9}$$

In this section, we prove a number of lemmas needed in the proof of Theorem 3.1. The first lemma states an elementary fact about the set $\mathcal{M}^+$.

Lemma 6.1. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Then

$$\mathcal{M}^+ \subseteq \{(b(x), x): 0 < x < L\} \tag{6.10}$$

i.e., if $(\beta, m) \in \mathcal{M}^+$, then $m \in (0, L)$ and $b(m) = \beta$.

Proof. According to part (d) of Theorem 2.2, if $(\beta, m) \in \mathcal{M}$, then $m \in (-L, L)$ and $\beta g'(m) = i'(m)$. If in addition $m > 0$, then $b(m) = \beta$. This yields the lemma. ∎

The next lemma considers the structure of the set $\mathcal{M}_\beta^+$ and defines a quantity β_1 that later is proved to be the smallest critical value (Lemma 6.6).

Lemma 6.2. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. The following conclusions hold:

(a) There exists $\beta > 0$ such that $\mathcal{M}_\beta^+ \neq \varnothing$.

(b) If $\mathcal{M}_\beta^+ \neq \varnothing$ and $\bar{\beta} > \beta$, then $0 \notin \mathcal{M}_{\bar{\beta}}$ and $\mathcal{M}_{\bar{\beta}}^+ \neq \varnothing$.

(c) The graph of $\mathcal{M}^+$ is strictly increasing; i.e., if (β, m) and $(\bar{\beta}, \bar{m})$ are in $\mathcal{M}^+$ with $\beta < \bar{\beta}$, then $m < \bar{m}$.

(d) Define the quantity

$$\beta_1 = \inf\{\beta > 0: \mathcal{M}_\beta^+ \neq \varnothing\}$$

Then $\beta_1 = \inf\{\beta > 0: 0 \notin \mathcal{M}_\beta\}$ and $\beta_1 > 0$.

(e) We also have $\beta_1 = \sup\{\beta > 0: \mathcal{M}_\beta^+ = \varnothing\} = \sup\{\beta > 0: \mathcal{M}_\beta = \{0\}\}$.

Proof. (a) For any fixed $x > 0$, we have $i(x) > 0$ and $g(x) > 0$. Hence, if we take $\beta > i(x)/g(x) > 0$, then

$$\beta g(x) - i(x) > 0 = \beta g(0) - i(0) \tag{6.11}$$

and so $0 \notin \mathcal{M}_\beta$. By part (d) of Theorem 2.2 and the evenness of g and i, there exists a number $m > 0$ such that

$$\beta g(m) - i(m) = \sup_{u \in \mathbb{R}}\{\beta g(u) - i(u)\} = -\beta\psi(\beta) \tag{6.12}$$

It follows that $\mathcal{M}_\beta^+ \neq \varnothing$.

(b) If $x \in \mathcal{M}_\beta^+$, then $x > 0$ and $g(x) > 0$ and for any $\bar{\beta} > \beta$

$$\bar{\beta} g(x) - i(x) > \beta g(x) - i(x) \geqslant \beta g(0) - i(0) = 0 \tag{6.13}$$

It follows that $0 \notin \mathscr{M}_\beta$. By part (d) of Theorem 2.2 and the evenness of g and i, we must have $\mathscr{M}_\beta^+ \neq \varnothing$.

(c) If $(\beta, m) \in \mathscr{M}^+$, then for all $z \in (0, L)$

$$0 \leqslant \beta g(m) - i(m) - [\beta g(z) - i(z)] = \int_z^m [\beta g'(y) - i'(y)]\, dy \qquad (6.14)$$

Now suppose that we have (β, m) and $(\bar{\beta}, \bar{m})$ in $\mathscr{M}^+$ with $\beta < \bar{\beta}$ but with $\bar{m} < m$. Then a twofold application of (6.14) would yield the contradiction

$$0 \leqslant \int_{\bar{m}}^m [\beta g'(y) - i'(y)]\, dy < \int_{\bar{m}}^m [\bar{\beta} g'(y) - i'(y)]\, dy$$

$$= -\int_m^{\bar{m}} [\bar{\beta} g'(y) - i'(y)]\, dy \leqslant 0 \qquad (6.15)$$

where we made use of (6.2). Hence, if (β, m) and $(\bar{\beta}, \bar{m})$ are in $\mathscr{M}^+$ with $\beta < \bar{\beta}$, then $m \leqslant \bar{m}$. On the other hand, if $\beta < \bar{\beta}$, but $m = \bar{m}$, then by Lemma 6.1

$$\beta = b(m) = b(\bar{m}) = \bar{\beta} \qquad (6.16)$$

and this contradicts $\beta < \bar{\beta}$. Hence, if $\beta < \bar{\beta}$, then $m < \bar{m}$. We have proved that the graph of $\mathscr{M}^+$ is strictly increasing.

(d) The quantity β_1 is finite by part (a). Define the quantity

$$A = \inf\{\beta > 0 \colon 0 \notin \mathscr{M}_\beta\}$$

If $0 \notin \mathscr{M}_\beta$, then $\mathscr{M}_\beta^+ \neq \varnothing$; hence, $\beta_1 \leqslant A$. If $\bar{\beta} > \beta_1$, then by part (b), $0 \notin \mathscr{M}_{\bar{\beta}}$ and so $\bar{\beta} \geqslant A$; hence $\beta_1 \geqslant A$. It follows that $\beta_1 = A$. Using the equality $\beta_1 = A$, we proved that $\beta_1 > 0$ in Lemma 4.3.

(e) Define the quantities

$$B = \sup\{\beta > 0 \colon \mathscr{M}_\beta^+ = \varnothing\}, \qquad C = \sup\{\beta > 0 \colon \mathscr{M}_\beta = \{0\}\} \qquad (6.17)$$

If $\beta > B$, then $\mathscr{M}_\beta^+ \neq \varnothing$ and so $\beta \geqslant \beta_1$; hence $B \geqslant \beta_1$. If $B > \beta_1$, then there exist β and $\bar{\beta}$ satisfying $B > \beta > \bar{\beta} > \beta_1$, $\mathscr{M}_\beta^+ = \varnothing$, $\mathscr{M}_{\bar{\beta}}^+ \neq \varnothing$. This contradiction to part (b) proves that $B = \beta_1$. Since $\mathscr{M}_\beta$ is nonempty for any $\beta > 0$ [Theorem 2.2(d)], it is clear that $\mathscr{M}_\beta^+ = \varnothing$ if and only if $\mathscr{M}_\beta = \{0\}$; hence, $\beta_1 = B = C$. This completes the proof of Lemma 6.2. ∎

Theorem 3.1 states the existence of the spontaneous magnetization $m(\beta)$ and states a number of its properties. The following lemma will be useful in analyzing this function. According to Lemma 6.2, there exists a positive number β_1 with the property that

$$\mathcal{M}_\beta^+ \begin{cases} = \varnothing & \text{for} \quad 0 < \beta < \beta_1 \\ \neq \varnothing & \text{for} \quad \beta > \beta_1 \end{cases} \tag{6.18}$$

For $\beta > \beta_1$, we define the quantities

$$m^-(\beta) = \sup\{m: m \in \mathcal{M}_{\beta'}^+, \beta_1 < \beta' < \beta\} \tag{6.19}$$

$$m^+(\beta) = \inf\{m: m \in \mathcal{M}_{\beta'}^+, \beta' > \beta\} \tag{6.20}$$

These quantities are well-defined because of (6.18).

Lemma 6.3. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. The following conclusions hold:

(a) For each $\beta > \beta_1$, $0 < m^-(\beta) \leqslant m^+(\beta) < L$.

(b) For each $\beta > \beta_1$, $(\beta, m^-(\beta))$ and $(\beta, m^+(\beta))$ are in $\mathcal{M}^+$. Thus

$$b(m^-(\beta)) = \beta = b(m^+(\beta)) \tag{6.21}$$

(c) If $m \in \mathcal{M}_\beta^+$ for some $\beta > \beta_1$, then $m^-(\beta) \leqslant m \leqslant m^+(\beta)$.

(d) Both $m^-(\beta)$ and $m^+(\beta)$ are strictly increasing functions of $\beta \in (\beta_1, \infty)$.

(e) $m^-(\beta)$ is a left-continuous function of $\beta \in (\beta_1, \infty)$ and $m^+(\beta)$ is a right-continuous function of $\beta \in (\beta_1, \infty)$.

Proof. (a) For any $\beta' \in (\beta_1, \beta)$, $\mathcal{M}_{\beta'}^+ \neq \varnothing$ [see (6.18)]. Hence, for any $m \in \mathcal{M}_{\beta'}^+$, we have $0 < m \leqslant m^-(\beta)$. For any $m \in \mathcal{M}_{\beta'}^+$, $\beta' \in (\beta_1, \beta)$, and any $\bar{m} \in \mathcal{M}_{\beta''}^+$, $\beta'' > \beta$, Lemma 6.2(d) guarantees that $m < \bar{m}$. It follows that $m^-(\beta) \leqslant m^+(\beta)$. For any $\bar{m} \in \mathcal{M}_{\beta'}$, $\beta' > \beta$, we have $m^+(\beta) \leqslant \bar{m} < L$ (Lemma 6.1). This completes the proof of part (a).

(b) We prove that $(\beta, m^-(\beta))$ is in $\mathcal{M}^+$. The proof that $(\beta, m^+(\beta))$ is in $\mathcal{M}^+$ is similar. For $n \in \{1, 2, \dots\}$ there exist numbers $\beta'_n \in (\beta_1, \beta)$ and $m_n \in \mathcal{M}_{\beta'_n}^+$ such that $m_n \uparrow m^-(\beta)$ as $n \to \infty$. By passing to a subsequence if necessary, we may assume that $\beta' = \lim_{n \to \infty} \beta'_n$ exists. Thus, $\beta_1 \leqslant \beta' \leqslant \beta$. For each n

$$\beta'_n g(m_n) - i(m_n) \geqslant \beta'_n g(x) - i(x) \qquad \text{for all} \quad x \in \mathbb{R} \tag{6.22}$$

Hence, by taking $n \to \infty$, we have

$$\beta' g(m^-(\beta)) - i(m^-(\beta)) \geqslant \beta' g(x) - i(x) \qquad \text{for all} \quad x \in \mathbb{R} \tag{6.23}$$

Since $m^-(\beta)>0$ [part (a)], it follows that $(\beta', m^-(\beta)) \in \mathcal{M}^+$. We now show that $\beta'=\beta$. If $\beta'<\beta$, then for any $\beta'' \in (\beta', \beta)$ and any $m \in \mathcal{M}_{\beta''}^+$, $m^-(\beta)<m$ [Lemma 6.2(c)]. This contradicts the definition (6.19) of $m^-(\beta)$ as a supremum and completes the proof of part (b).

(c) As in the proof of part (b), for $n \in \{1, 2,...\}$ there exist numbers $\beta'_n \in (\beta_1, \beta)$ and $m_n \in \mathcal{M}_{\beta'_n}^+$ such that $\beta'_n \to \beta$ and $m_n \to m^-(\beta)$ as $n \to \infty$. If $m \in \mathcal{M}_\beta^+$, then by Lemma 6.2(c), $m_n < m$. It follows that $m^-(\beta) \leqslant m$. A similar proof shows that if $m \in \mathcal{M}_\beta^+$, then $m \leqslant m^+(\beta)$. The equalities in (6.21) follow from Lemma 6.1.

(d) For $\beta_1<\beta<\beta'$, $m^-(\beta) \in \mathcal{M}_\beta^+$ and $m^-(\beta') \in \mathcal{M}_{\beta'}^+$ [part (b)], and so $m^-(\beta)<m^-(\beta')$ [Lemma 6.2(c)]. By a similar proof, $m^+(\beta)<m^+(\beta')$ for $\beta_1<\beta<\beta'$.

(e) We prove that if $\beta_k \uparrow \beta$, $\beta_k \in (\beta_1, \beta)$, then $m^-(\beta_k) \to m^-(\beta)$. The proof that $m^+(\beta)$ is a right-continuous function of $\beta \in (\beta_1, \infty)$ is similar. By Lemma 6.2(c), $m^-(\beta_k)<m^-(\beta)$, and so

$$\limsup_{k \to \infty} m^-(\beta_k) \leqslant m^-(\beta) \tag{6.24}$$

As in the proof of part (b), for $n \in \{1, 2,...\}$ there exist numbers $\beta'_n \in (\beta_1, \beta)$ and $m_n \in \mathcal{M}_{\beta'_n}^+$ such that $\beta'_n \to \beta$ and $m_n \uparrow m^-(\beta)$ as $n \to \infty$. Given $\varepsilon>0$, there exists an n such that $m_n \geqslant m^-(\beta)-\varepsilon$. Since for all sufficiently large k, $\beta_k > \beta'_n$, we have

$$m^-(\beta_k) > m_n \geqslant m^-(\beta)-\varepsilon \tag{6.25}$$

for all sufficiently large k [Lemma 6.2(c)]. It follows that

$$\liminf_{k \to \infty} m^-(\beta_k) \geqslant m^-(\beta)-\varepsilon \tag{6.26}$$

Taking $\varepsilon \to 0$ and using (6.24), we conclude that $m^-(\beta)$ is a left-continuous function of $\beta \in (\beta_1, \infty)$. This completes the proof of the lemma. ∎

Let us extend the definitions of $m^-(\beta)$ and $m^+(\beta)$ to $\beta \in (0, \beta_1]$ by setting

$$m^-(\beta)=m^+(\beta)=0 \qquad \text{for} \quad \beta \in (0, \beta_1) \tag{6.27}$$

and

$$m^-(\beta_1)=0, \qquad m^+(\beta_1)=\lim_{\beta \downarrow \beta_1} m^+(\beta) \tag{6.28}$$

The extended function $m^-(\beta)$ is a left-continuous function of $\beta \in (0, \infty)$, while the extended function $m^+(\beta)$ is a right-continuous function of $\beta \in (0, \infty)$ [see Lemma 6.3(e)].

The following lemma will be useful in analyzing the behavior of the spontaneous magnetization in a neighborhood of β_1 in the case $m^+(\beta_1)=0$. The fact stated here was proved in the course of proving part (c) of Theorem 4.4.

Lemma 6.4. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. If $m^+(\beta_1)=0$, then the function b can be extended to a real analytic function on $(-L, L)$.

In Definition 4.1, we defined the concepts of a first-order phase transition and a second-order phase transition. We now restate the definition, but for convenience we give condition (4.2) in an equivalent form $[m^-(\beta)=m^+(\beta)=\bar{m}$ and $b'(\bar{m})=0]$. The equivalence between (4.2) and (6.30) can be easily shown as in (7.7)–(7.8) below.

Definition 6.5. The generalized Curie–Weiss model has a *phase transition* at an inverse temperature $\beta \in [\beta_1, \infty)$ if either

$$m^-(\beta)<m^+(\beta) \tag{6.29}$$

(*first-order phase transition*) or

$$m^-(\beta)=m^+(\beta)=\bar{m} \qquad \text{and} \qquad b'(\bar{m})=0 \tag{6.30}$$

(*second-order phase transition*). Any value β at which a phase transition occurs is called a *critical value.*

The next lemma states some facts about phase transitions.

Lemma 6.6. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. Let $\mathscr{P}$ denote the set of all inverse temperatures at which the generalized Curie–Weiss model has a phase transition. The following conclusions hold.

(a) $\mathscr{P} \neq \varnothing$. In fact, the quantity β_1 with the property given in (6.18) belongs to $\mathscr{P}$ and is the smallest element of $\mathscr{P}$.

(b) For any $\beta \in \mathscr{P}$ there exists a value $\bar{m} \in [m^-(\beta), m^+(\beta)]$ such that

$$b'(\bar{m})=0 \tag{6.31}$$

(c) $\mathscr{P}$ is locally finite; i.e., for any $\gamma>0$ the set $\mathscr{P} \cap [0, \gamma]$ is finite.

Proof. (a) Clearly, $m^+(\beta_1)=\lim_{\beta \downarrow \beta_1} m^+(\beta) \geqslant 0$. If $m^+(\beta_1)>0= m^-(\beta)$, then we have a first-order phase transition at β_1. Suppose that $m^+(\beta_1)=0=m^-(\beta_1)$. According to Lemma 6.4, b can be extended to a real analytic function on $(-L, L)$, and

$$b'(m^+(\beta))=b'(0)=0 \tag{6.32}$$

[see (6.5)]. Thus, we have a second-order phase transition at β_1. We conclude that whether $m^+(\beta_1) > 0$ or $m^+(\beta_1) = 0$, β_1 belongs to $\mathscr{P}$. The fact that $\beta_1 = \min\{\beta: \beta \in \mathscr{P}\}$ is now a consequence of (6.18).

(b) If there is a second-order phase transition at β, then by definition $b'(\bar{m}) = 0$, where $\bar{m} = m^-(\beta) = m^+(\beta)$. If there is a first-order phase transition at β, then $m^-(\beta) < m^+(\beta)$ and by (6.21)

$$b(m^-(\beta)) = \beta = b(m^+(\beta)) \tag{6.33}$$

By the mean value theorem, there exists $\bar{m} \in [m^-(\beta), m^+(\beta)]$ such that $b'(\bar{m}) = 0$.

(c) According to part (a), $\beta_1 = \min\{\beta: \beta \in \mathscr{P}\}$. It suffices to prove that for any $\gamma > \beta_1$ the set $(\mathscr{P} \setminus \{\beta_1\}) \cap [0, \gamma]$ is finite. According to part (b), there exists a mapping from $(\mathscr{P} \setminus \{\beta_1\}) \cap [0, \gamma]$ into $[0, L)$ that associates to each $\beta \in \mathscr{P}$ a root of b' in the interval $[m^-(\beta), m^+(\beta)]$. This mapping is injective since $\beta_1 < \beta < \bar{\beta}$ implies that $m^+(\beta) < m^-(\bar{\beta})$ [Lemmas 6.3(b) and 6.2(c)]. We claim that the image of $(\mathscr{P} \setminus \{\beta_1\}) \cap [0, \gamma]$ under this mapping is a subset of the interval $[m^+(\beta_1), m^+(\gamma)]$. Indeed, if $\bar{\beta} > \beta_1$, then for any $\beta \in (\beta_1, \bar{\beta})$, $m^+(\beta) < m^-(\bar{\beta})$, and so

$$m^+(\beta_1) = \lim_{\beta \downarrow \beta_1} m^+(\beta) \leqslant m^-(\bar{\beta}) \tag{6.34}$$

Thus, if $\beta_1 < \bar{\beta} \leqslant \gamma$, then $m^+(\beta_1) \leqslant m^-(\bar{\beta}) \leqslant m^+(\bar{\beta}) \leqslant m^+(\gamma)$. This proves the claim. It follows that the set $(\mathscr{P} \setminus \{\beta_1\}) \cap [0, \gamma]$ is injectively mapped into the set of roots of b' in the interval $[m^+(\beta_1), m^+(\gamma)]$. Now either $m^+(\beta_1) > 0$ or $m^+(\beta_1) = 0$. If $m^+(\beta_1) > 0$, then the functions b and b', being real analytic on $(-L, 0) \cup (0, L)$, are real analytic in a neighborhood of the interval $[m^+(\beta_1), m^+(\gamma)]$. Hence, there are only finitely many roots of b' in this interval. This shows that the set $(\mathscr{P} \setminus \{\beta_1\}) \cup [0, \gamma]$ is finite. If $m^+(\beta_1) = 0$, then by Lemma 6.4, b and thus b' can be extended to a real analytic function on $(-L, L)$. Again, we conclude that there are only finitely many roots of b' in the interval $[m^+(\beta_1), m^+(\gamma)]$ and therefore that the set $(\mathscr{P} \setminus \{\beta_1\}) \cap [0, \gamma]$ is finite. This completes the proof of the lemma. ▮

The next lemma, which is the last lemma in this section, is needed in the proof of part (f) of Theorem 3.1.

Lemma 6.7. We assume that the interaction function g and the single-site distribution ρ satisfy Hypotheses 1.1. If $L < \infty$ and $\rho\{L\} > 0$, then

$$\lim_{t \to \infty} tc''(t) = 0 \tag{6.35}$$

where $c(t) = \log \int_{\mathbb{R}} \exp(tx)\, \rho(dx)$ for t real.

Proof. Let $\alpha = \rho\{L\} > 0$ and define for t real the probability measure

$$\rho_t(dx) = \exp(tx)\,\rho(dx) \Big/ \int_{\mathbb{R}} \exp(ty)\,\rho(dy) \tag{6.36}$$

Then

$$c''(t) = \int_{\mathbb{R}} x^2\,\rho_t(dx) - \left[\int_{\mathbb{R}} x\,\rho_t(dx)\right]^2 > 0 \tag{6.37}$$

Since the support of ρ_t is a subset of the closed interval $[-L, L]$, we conclude that for $t > 0$

$$\begin{aligned} 0 \leqslant tc''(t) &\leqslant t\left\{\int Lx\,\rho_t(dx) - \left[\int x\,\rho_t(dx)\right]^2\right\} \\ &= \left[\int x\,\rho_t(dx)\right] \int t(L-x)\,\rho_t(dx) \\ &\leqslant L \int t(L-x)\exp(tx)\,\rho(dx)/[\alpha\exp(tL)] \\ &= (L/\alpha)\int_{[-L,L]} t(L-x)\exp[-t(L-x)]\,\rho(dx) \end{aligned} \tag{6.38}$$

For each $t > 0$ the nonnegative functions $f_t(x) = t(L-x)\exp[-t(L-x)]$, $x \in [-L, L]$, are bounded by $1/e$ and as $t \to \infty$, $f_t(x) \to 0$ for each $x \in [-L, L]$. Hence, by the dominated convergence theorem

$$\int_{[-L,L]} f_t(x)\,\rho(dx) \to 0 \qquad \text{as} \quad t \to \infty \tag{6.39}$$

It follows from (6.38) that $tc''(t) \to 0$ as $t \to \infty$. ∎

In the next section, we will use the above lemmas in order to complete the proof of Theorem 3.1.

7. PROOF OF THEOREM 3.1

Using the lemmas presented in the previous section, we now prove Theorem 3.1.

7.1. Proof of Theorem 3.1(a)

The set $\mathscr{P}$ denotes the set of all inverse temperatures at which the generalized Curie–Weiss model has a phase transition. Each $\beta \in \mathscr{P}$ is called a critical value. According to Lemma 6.6, the set $\mathscr{P}$ is nonempty and is locally finite. Thus, the elements of $\mathscr{P}$ are either finite in number ($0<\beta_1<\cdots<\beta_N$ for some $N\in\{1,2,\ldots\}$) or countably infinite ($0<\beta_1<\beta_2<\cdots$) and divergent to ∞.

For $\beta\in(0,\infty)\backslash\mathscr{P}$, we define

$$m(\beta)=m^-(\beta)=m^+(\beta) \tag{7.1}$$

By (6.27), $m(\beta)=0$ for $\beta\in(0,\beta_1)$, while by Lemma 6.3, $m(\beta)$ is positive and is strictly increasing for $\beta\in(\beta_1,\infty)\backslash\mathscr{P}$. Since $m^+(\beta)$ is right continuous and $m^-(\beta)$ is left continuous, $m(\beta)$ is continuous on the connected open subintervals of the set $(0,\infty)\backslash\mathscr{P}$. According to Lemma 6.3(b),

$$b(m(\beta))=\beta \qquad \text{for} \quad \beta\in(\beta_1,\infty)\backslash\mathscr{P} \tag{7.2}$$

Since we cannot have a second-order phase transition at any $\beta\notin\mathscr{P}$, it follows from (6.30) that

$$b'(m(\beta))\neq 0 \qquad \text{for} \quad \beta\in(\beta_1,\infty)\backslash\mathscr{P} \tag{7.3}$$

Putting the information together, we see that on the connected open subintervals of the set $(\beta_1,\infty)\backslash\mathscr{P}$, $m(\beta)$ is a continuous, strictly increasing function whose inverse function is the real analytic function b. The inverse function theorem implies that $m(\beta)$ is real analytic on the connected open subintervals of the set $(\beta_1,\infty)\backslash\mathscr{P}$. Of course, for $\beta\in(0,\beta_1)$, $m(\beta)\equiv 0$ is real analytic.

According to the Definition 6.5 of phase transition, there are two possibilities for the behavior of $m(\beta)$ in a neighborhood of a critical value $\beta_i\in\mathscr{P}$. The first possibility is that by (6.29), m has a jump discontinuity at β_i:

$$\lim_{\beta\uparrow\beta_i} m(\beta)=m^-(\beta_i)<m^+(\beta_i)=\lim_{\beta\downarrow\beta_i} m(\beta) \tag{7.4}$$

The second possibility is that by (6.30)

$$m^-(\beta_i)=m^+(\beta_i) \qquad \text{and} \qquad b'(m^+(\beta_i))=0 \tag{7.5}$$

If we define $m(\beta_i)=m^+(\beta_i)$, then $m(\beta)$ is continuous at β_i. In this second case, we would like to show that

$$\lim_{\beta\downarrow\beta_i} m'(\beta)=\infty \tag{7.6}$$

Since $b(m(\beta)) = \beta$ for $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$ [see (7.2)], we have for $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$

$$m'(\beta) = \frac{1}{b'(m(\beta))} > 0 \tag{7.7}$$

For $i \geqslant 2$, $m^+(\beta_i) \in (0, L)$, and so b' is real analytic in a neighborhood of the point $m^+(\beta_i)$. Hence

$$\lim_{\beta \downarrow \beta_i} m'(\beta) = \lim_{\beta \downarrow \beta_i} \frac{1}{b'(m(\beta))} = \frac{1}{b'(m^+(\beta_i))} = \infty \tag{7.8}$$

For $i = 1$, $m^+(\beta_1) = 0$, in which case b can be extended to a real analytic function on $(-L, L)$. By the same calculation of (7.8) with $i = 1$, we see that again $\lim_{\beta \downarrow \beta_1} m'(\beta) = \infty$. Because of the two possibilities for the behavior of $m(\beta)$ in the neighborhood of a critical value $\beta_i \in \mathscr{P}$, $m(\beta)$ cannot be represented as the restriction of one real analytic function in a neighborhood of a critical value $\beta_i \in \mathscr{P}$. This completes the proof of part (a) of Theorem 3.1.

7.2. Proof of Theorem 3.1(b)

According to (6.18), the critical value β_1 has the property that $\mathscr{M}_\beta^+ = \varnothing$ for $\beta \in (0, \beta_1)$. Since $\mathscr{M}_\beta$ is nonempty for all $\beta > 0$ [Theorem 2.2(d)], it follows that for $0 < \beta < \beta_1$, $\mathscr{M}_\beta = \{0\}$. This gives the first line of part (b). According to part (c) of Lemma 6.3, for all $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$ the set $\mathscr{M}_\beta^+$ consists of the single point $\{m(\beta)\}$. Hence, by symmetry, for all $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$ the set $\mathscr{M}_\beta$ equals $\{m(\beta), -m(\beta)\}$. This gives the second line of part (b).

7.3. Proof of Theorem 3.1(c)

This is proved in parts (d) and (e) of Lemma 6.2.

7.4. Proof of Theorem 3.1(d)

For $\beta \in (0, \beta_1)$, $m(\beta) = 0$, and thus (3.11) holds:

$$\beta g'(m(\beta)) = 0 = i'(m(\beta)) \tag{7.9}$$

For $\beta \in (\beta_1, \infty) \backslash \mathscr{P}$, we have

$$b(m(\beta)) = \frac{i'(m(\beta))}{g'(m(\beta))} = \beta \tag{7.10}$$

where $g'(m(\beta)) > 0$, since $m(\beta) > 0$. Thus, again (3.11) holds.

We now prove that

$$m(\beta) \uparrow L \qquad \text{as} \quad \beta \to \infty \quad \text{in the set } (0, \infty) \backslash \mathscr{P} \tag{7.11}$$

Suppose that (7.11) fails. Then there exists some $r \in (0, L)$ and a sequence $\{\beta_N, N = 1, 2,...\}$ satisfying $\beta_N \in (\beta_1, \infty) \backslash \mathscr{P}$, $\beta_N \to \infty$, such that $m(\beta_N) \leqslant r$ for each N. Since $\beta_N > \beta_1$ for each N, we have $m(\beta_N) > m^+(\beta_1)$. Below we will prove that the image of the interval $[m^+(\beta_1), r]$ is a bounded subset of $\mathbb{R}$; i.e., that

$$b([m^+(\beta_1), r]) \subseteq [c, d], \qquad \text{some} \quad -\infty < c < d < \infty \tag{7.12}$$

Then for each N, since $b(m(\beta_N)) = \beta_N$ and $m^+(\beta_1) < m(\beta_N) \leqslant r$, we would have

$$b(m(\beta_N)) = \beta_N \leqslant d \tag{7.13}$$

But (7.13) is impossible, since $\beta_N \to \infty$. This contradiction proves (7.11). We now prove (7.12). If $m^+(\beta_1) > 0$, then the function b, being real analytic on $(-L, 0) \cup (0, L)$, is real analytic in a neighborhood of the interval $[m^+(\beta_1), r]$. In this case, (7.12) clearly holds. If $m^+(\beta_1) = 0$, then by Lemma 6.4, b can be extended to a real analytic function on $(-L, L)$, and again (7.12) holds. This completes the proof of part (d).

7.5. Proof of Theorem 3.1(e)

Because of the structure of $\mathscr{M}_\beta$ given in part (b), the limit (3.13) is proved exactly like the analogous limit for the Curie–Weiss model as presented in Section IV.4 of ref. 4 (see Theorem IV.4.1 there).

7.6. Proof of Theorem 3.1(f)

We assume that $L < \infty$ and $\rho\{L\} > 0$. If $\mathscr{P}$ consists of the single critical point $\{\beta_1\}$, then we are done, so let us assume that the cardinality of $\mathscr{P}$ is at least two. According to part (b) of Lemma 6.6, there exists a mapping from $\mathscr{P} \backslash \{\beta_1\}$ into $[0, L)$ that associates to each $\beta_i \in \mathscr{P} \backslash \{\beta_1\}$ a root of b' in the interval $[m^-(\beta_i), m^+(\beta_i)]$. This mapping is injective since $\beta_i < \beta_j$, $1 < i < j$, implies that $m^+(\beta_i) < m^-(\beta_j)$ [Lemmas 6.3(b) and 6.2(c)]. It follows that the set $\mathscr{P} \backslash \{\beta_1\}$ is injectively mapped into the set of roots of b' in the interval $[m^-(\beta_2), L)$. The function b is real analytic in this interval. If we prove that $b'(x) \to \infty$ as $x \uparrow L$, then there can be only finitely many roots of b' in the interval $[m^-(\beta_2), L)$, and part (f) will follow.

By Property 2.1(i) of $i(u)$, $i'(c'(t)) = t$ for all $t \in \mathbb{R}$. Thus,

$$\frac{i''(u)}{i'(u)} = \frac{1}{(c')^{-1}(u)\, c''((c')^{-1}(u))} \tag{7.14}$$

As $u \uparrow L$, $(c')^{-1}(u) \to \infty$, and so Lemma 6.7 implies that $i''(u)/i'(u) \to \infty$ as $u \uparrow L$. In a neighborhood of L, the functions g' ad $1/g'$ are positive and bounded and g'' is bounded. Since $i'(u) \to \infty$ as $u \uparrow L$, we find

$$\begin{aligned} b'(u) &= i''(u)/g'(u) - i'(u)\, g''(u)/[g'(u)]^2 \\ &= i'(u)\{i''(u)/[i'(u)\, g'(u)] - g''(u)/[g'(u)]^2\} \to \infty \end{aligned} \tag{7.15}$$

as $u \uparrow L$. This completes the proof of part (f) of Theorem 3.1. With part (f), the proof of Theorem 3.1 is finished.

ACKNOWLEDGMENTS

One of us (R.S.E.) would like to thank Jonathan Machta for several useful discussions and for pointing out the paper of Mouritsen, Frank, and Mukamel (1983).

REFERENCES

1. M. Cassandro, E. Olivieri, A. Pellegrinotti, and E. Presutti, *Z. Wahrsch. Verw. Geb.* **41**:313–334 (1978).
2. F. J. Dyson, in *Mathematical Aspects of Statistical Mechanics*, J. C. T. Pool, ed. (American Mathematical Society, Providence, Rhode Island, 1972), pp. 1–11.
3. T. Eisele and R. S. Ellis, *Z. Wahrsch. Verw. Geb.* **63**:297–348 (1983).
4. R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics* (Springer, New York, 1985).
5. R. S. Ellis, J. L. Monroe, and C. M. Newman, *Commun. Math. Phys.* **46**:167–182 (1976).
6. R. S. Ellis and C. M. Newman, *Z. Wahrsch. Verw. Geb.* **44**:117–139 (1978).
7. R. S. Ellis and C. M. Newman, *Trans. Am. Math. Soc.* **237**:83–99 (1978).
8. R. S. Ellis, C. M. Newman, and J. S. Rosen, *Z. Wahrsch. Verw. Geb.* **51**:153–169 (1980).
9. M. E. Fisher, *Phys. Rev.* **162**:480–485 (1967).
10. R. B. Griffiths, *Commun. Math. Phys.* **6**:121–127 (1967).
11. K. Husimi, in *Proceedings of the International Conference of Theoretical Physics*, H. Yukawa, ed. (Science Council of Japan, Tokyo, 1953), pp. 531–533.
12. J.-W. Jeon, Central limit theorems in the regions of large deviations with applications to statistical mechanics, Dissertation, Florida State University (1979).
13. J.-W. Jeon, in *Proceedings of the 3rd Japan–Korea Joint Conference of Statistics* (Kyoto University, 1985).
14. M. Kac, in *Statistical Physics: Phase Transitions and Superfluidity*, M. Chretien, E. P. Gross, and S. Deser (Gordon and Breach, New York, 1966), Vol. 1, pp. 241–305.
15. O. G. Mouritsen, B. Frank, and D. Mukamel, *Phys. Rev. B* **27**:3018–3031 (1983).

16. C. M. Newman, Shock waves and mean field bounds—Concavity and analyticity of the magnetization at low temperature, Preprint, University of Arizona; also, talk given at 46th Statistical Mechanics Meeting, Rutgers University (December 1981).
17. P. A. Pearce, *J. Stat. Phys.* **25**:309–320 (1981).
18. B. Simon, *J. Stat. Phys.* **22**:491–493 (1980).
19. J. Slawny, *J. Stat. Phys.* **32**:375–388 (1983).
20. A. D. Sokal, *J. Stat. Phys.* **28**:431–439 (1982).
21. H. N. V. Temperley, *Proc. Phys. Soc. A* **67**:233–238 (1954).
22. C. J. Thompson, *Commun. Math. Phys.* **24**:61–66 (1971).
23. C. J. Thompson, *Mathematical Statistical Mechanics* (Macmillan, New York, 1972).
24. J. O. Vigfusson, *Lett. Math. Phys.* **10**:71–77 (1985).