Journal of Statistical Physics, Vol. 47, Nos. 3/4, 1987

The Phase Transition in a General Class of Ising-Type Models is Sharp

M. Aizenman,[1] D. J. Barsky,[2] and R. Fernández[3]

Received December 23, 1986

For a family of translation-invariant, ferromagnetic, one-component spin systems—which includes Ising and φ^4 models—we prove that (i) the phase transition is sharp in the sense that at zero magnetic field the high- and low-temperature phases extend up to a common critical point, and (ii) the critical exponent $\hat{\beta}$ obeys the mean field bound $\hat{\beta} \leqslant 1/2$. The present derivation of these nonperturbative statements is not restricted to "regular" systems, and is based on a new differential inequality whose Ising model version is $M \leqslant \beta h \chi + M^3 + \beta M^2 \, \partial M/\partial \beta$. The significance of the inequality was recognized in a recent work on related problems for percolation models, while the inequality itself is related to previous results, by a number of authors, on ferromagnetic and percolation models.

KEY WORDS: Phase transition; Ising model; φ^4; intermediate phase; critical exponents; inequalities.

1. INTRODUCTION

In this paper we show that for a general class of one-component ferromagnetic spin models the phase transition occurs directly from the high-temperature regime characterized by exponential decay of correlation functions to the low-temperature regime where there is spontaneous magnetization. This statement, and a related mean field bound on the critical exponent $\hat{\beta}$, are proven here for translation-invariant, and even just periodic, interactions of any range. The problem is of interest since it concerns a basic issue, the resolution of which requires analysis of a nonpertur-

[1] Departments of Mathematics and Physics, Rutgers University, New Brunswick, New Jersey 08903.

[2] Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903.

[3] Department of Mathematics, University of Texas at Austin, Austin, Texas 78712.

0022-4715/87/0500-0343$05.00/0

bative regime. Previously such results were proven only under a regularity assumption, which has been established only for a restricted collection of models[1] (which includes, nevertheless, the nearest neighbor case in dimensions $d > 2$).

Our work is based on the analysis of the analogous percolation problem contained in Ref. 2, where a technique was developed for dealing with this problem by means of a new differential inequality in the relevant two-parameter space. An analogous inequality is proven here for Ising models and other ferromagnetic systems with spins in the Griffiths–Simon class, making the analysis of Ref. 2 directly applicable also to them. Although there are many similarities between the two relations, their proofs require different techniques. One also learns much about the two models from their differences: the behavior of Ising models reflects an underlying connection with φ^4 fields, while percolation is related to φ^3 field theory. The nonlinear partial differential inequality derived here is also of independent interest, and we shall comment below on its relations with the results of Aizenman and Graham,[3] Fröhlich and Sokal,[4] Chayes and Chayes,[5] and Aizenman and Barsky.[2]

We consider here systems of one-component spin variables on the d-dimensional lattice $\mathbb{Z}^d$, whose Hamiltonians are the sum of ferromagnetic pair interactions:

$$H(\sigma) = -(1/2) \sum_{x,y} J_{x-y} \sigma_x \sigma_y - h \sum_x \sigma_x \tag{1.1}$$

with $J_z \geqslant 0$. The specification of $\mathbb{Z}^d$ as the lattice is only for the sake of concreteness; the key feature of the model that will be used is the translation invariance, which is manifest in the Hamiltonian (1.1). Furthermore, our methods may also be extended to a class of weakly inhomogeneous systems, which includes the periodic models. It will be assumed that $|J| = \sum_x J_x < \infty$, since otherwise the spins are totally ordered at any finite temperature.

The equilibrium states of such systems are described by probability measures whose formal expression is

$$\rho(d\sigma) = \rho_0(d\sigma)\, e^{-\beta H(\sigma)}/\text{norm}$$

where β is the inverse temperature and ρ_0 is the product of the *a priori* single-site measures describing noninteracting spins. The analysis presented here applies to spin variables in the *Griffiths–Simon (GS) class*, whose distributions allow the spin σ to be written either as a block variable for a system of ferromagnetically coupled Ising spins or as a distributional limit of such variables. (See Section 4 and Refs. 3, 6, and 7.) Two important examples of such spin variables are the Ising spins themselves and the "φ^4 variables."

For such spin systems the Lee–Yang[8] theory applies and shows that phase transitions may occur only at $h = 0$, where a phase transition is known to exist if the dimensionality exceeds 1,[9] or if $d = 1$ and the interactions decay no faster than $c/|x - y|^2$.[10,11] The phase transition is manifested by symmetry-breaking, and the nonvanishing of the spontaneous (or residual) magnetization

$$M_+(\beta) \equiv \lim_{h \to 0^+} \langle \sigma_0 \rangle_{\beta,h}$$

where $\langle \cdot \rangle_{\beta,h}$ denotes the expectation value in an infinite-volume Gibbs state.

Our purpose here is to give a nonperturbative analysis of the phase structure observed in these models by varying β, with h fixed at zero. We prove that along this line in the parameter space, the high-temperature phase and the low-temperature phase actually extend up to a common critical point. While the two phases are relatively well understood in the corresponding high- and low-temperature regimes, the result described here refers to a region that is not accessible to direct expansion methods. An example of the possibility excluded here for one-component models is provided by the Kosterlitz–Thouless phase of two-component spin systems in two dimensions with a Hamiltonian similar to (1.1), in which the system exhibits neither long-range order nor rapid decay of correlations.[12,13] For any translation-invariant ferromagnetic model with one-component spins (in the GS class) that behavior is proven here to be limited to only a single critical point.

Following is the precise statement of our result.

Theorem 1. For Ising, φ^4, and other variables in the Griffiths–Simon class with the Hamiltonian (1.1) on $\mathbb{Z}^d$ ($d \geqslant 1$), there is a $\beta_c \in [0, \infty]$ with the following properties.

(i) For all $0 \leqslant \beta < \beta_c$ there is a unique infinite-volume Gibbs state, the spontaneous magnetization vanishes [i.e., $M_+(\beta) = 0$], and the magnetic susceptibility is finite:

$$\chi(\beta, 0) \equiv \sum_{x \in \mathbb{Z}^d} \langle \sigma_0 \sigma_x \rangle_{\beta, h=0} < \infty \tag{1.2}$$

(ii) For all $\beta > \beta_c$ there is symmetry-breaking, with

$$M_+(\beta) \geqslant \text{const} \cdot [(\beta - \beta_c)/\beta_c]^{1/2} \tag{1.3}$$

(iii) At $\beta = \beta_c$, the magnetization in a positive magnetic field $h \to 0^+$ decays no faster than the following bound:

$$M(\beta_c, h) \equiv \langle \sigma_0 \rangle_{\beta_c, h} \geqslant \text{const} \cdot h^{1/3} \tag{1.4}$$

It should be noted that the last bound was first derived by Fröhlich and Sokal[4] for a slightly different class of spin distributions, which also includes the Ising and the "φ^4 variables." The inequality (1.4) is included in the above statement because it forms for us an integral part of the wider picture, and because it is proven here at the same level of generality as (i) and (ii).

Remarks. (i) The finiteness of the susceptibility (1.2) is known to imply rapid decay of correlations; see Refs. 1 and 13.

(ii) The bound (1.3) implies also, by the arguments of Lebowitz,[15] that for any $\beta > \beta_c$ the state constructed with free boundary conditions exhibits long-range order.

(iii) The inequalities (1.3) and (1.4) imply of course bounds on critical exponents. Specifically, for the exponents $\hat{\beta}$ and δ defined as

$$\hat{\beta} = \limsup_{\beta \to \beta_c +} \log M_+(\beta)/\log(\beta - \beta_c) \tag{1.5}$$

and

$$\delta = \liminf_{h \to 0^+} \log h/\log M(\beta_c, h) \tag{1.6}$$

(1.3) and (1.4) yield

$$\hat{\beta} \leqslant 1/2 \tag{1.7}$$

and

$$\delta \geqslant 3 \tag{1.8}$$

(assuming that the critical temperature is finite, i.e., $\beta_c < \infty$). It might be noted here that 1/2 and 3 are the mean field values of $\hat{\beta}$ and δ, which are in fact attained in high dimensions for the Ising model (as finally proven in Ref. 16).

The above general extend those previously obtained in Ref. 1, where the absence of an intermediate regime—and the bound (1.3)—were proven under a certain additional "regularity hypothesis." It might be pointed out that the analysis of Ref. 1 does apply to the nearest neighbor models in dimensions $d > 2$ and to various other reflection positive models. The regularity condition introduced there was used more explicitly for other purposes, for which it may indeed be essential. However, as we show here, regularity is not relevant for the basic properties described by Theorem 1. That fact was expected, since unlike for the other results derived in Ref. 1, the function $f(\chi)$ that enters in the regularity hypothesis does not appear in the bound (1.3).

The proof of Theorem 1 rests on the following new nonlinear differential inequality (which for Ising spins has a version slightly simpler than the general one), and on the analysis of Ref. 2, where the significance of such a relation was discovered in the context of independent percolation models. In presenting the inequality we restrict our attention to finite systems; specifically, the squares $(-L, L]^d$ with periodic interactions (e.g., $J_x^{(L)} = \sum_{z \in \mathbb{Z}^d} J_{x+2Lz}$).

Theorem 2. In a ferromagnetic spin model in $(-L, L]^d$ with the Hamiltonian (1.1), with $J_{x,y}$ invariant under cyclic shifts, and spins having as the *a priori* distribution a fixed measure in the Griffiths–Simon class, the magnetization $M(\beta, h) = \langle \sigma_0 \rangle_{\beta,h}$ obeys the following bounds:

(i) If $\{\sigma_x\}$ are all Ising spins, then

$$M \leqslant \tanh(\beta h)\, \partial M/\partial(\beta h) + M^2(\beta\, \partial M/\partial\beta + M) \tag{1.9}$$

(ii) In the more general case

$$M \leqslant (\beta h)\, \partial M/\partial(\beta h) + (\beta\, |J|\, M^2 + \beta h M)(\beta\, \partial M/\partial\beta + M) \tag{1.10}$$

Let us remark that in this work we choose the relevant parameters of the phase diagram to be β and βh. In particular, $\partial M/\partial(\beta h)$ and $\partial M/\partial\beta$ denote here derivatives performed at constant β and constant βh, respectively. However, we shall continue to write $M = M(\beta, h)$.

It is instructive to compare (1.9) and (1.10) with the inequality

$$M \leqslant h\, \partial M/\partial h + M(\beta\, \partial M/\partial\beta + M) \tag{1.11}$$

derived Ref. 2 for the order parameter in independent bond percolation models whose bonds are occupied with the densities $K_{x,y} = 1 - \exp(-\beta J_{x-y})$. The reader is referred to Ref. 2 for the precise definition of the function M in that context. Here let us just say that when $h = 0$, $M(\beta, h)$ reduces to the percolation density, i.e., the probability that a given site belongs to an infinite cluster. The inequalities (1.9) and (1.10) differ from (1.11) essentially only in the powers of M in the last terms. A cursory inspection of the derivations of these inequalities permits tracing this difference to a revealing statement about the two families of models. The graphical structures that appear naturally in the analysis of percolation models (see Fig. 2 in Ref. 2) resemble those found in a diagrammatic expansion of a φ^3 field theory, whereas the graphs arising in the study of models in the Griffiths–Simon class are those of a φ^4 field theory (cf. Fig. 2 below). Such comparisons have been noted before, both on a nonrigorous level, where the critical behavior of the discrete models was analyzed (without full justification) assuming the applicability of field-theoretic methods,[17,18] and on the rigorous level, where the models were studied directly without any such assumptions.[19,20]

In addition to inequality (1.11), the analysis of Ref. 2 also required the inequality

$$\partial M/\partial\beta \leqslant \beta\,|J|\,M\,\partial M/\partial(\beta h) \tag{1.12}$$

as well as the observation (trivial in percolation models)

$$\partial M/\partial(\beta h) \leqslant M/\beta h \tag{1.13}$$

For the models considered in this paper, inequalities (1.12) and (1.13) are both consequences of the Griffiths–Hurst–Sherman inequality.[21]

It may seem somewhat surprising that relations (1.11)–(1.13) (with some auxiliary information which also applies to both models) imply the result on the sharpness of the phase transition. That, however, is the main result of Ref. 2. In fact, since the inequalities (1.9) and (1.10) are stronger than their percolation counterpart (1.11), part (i) of Theorem 1 requires no further arguments here, once the inequalities are derived. The bounds claimed in parts (ii) and (iii) are stronger than what is generally valid for percolation. However, the analysis of Ref. 2 was formulated so as to yield the proper results from the stronger inequalities that we have here.

Let us mention here that the arguments used to prove Theorem 1 from (1.9)—or (1.10) in the more general setting—involve the reduction of each these partial differential inequalities to a pair of ordinary differential inequalities (ODIs). In particular, one of these ODIs is quite similar to the following bound of Fröhlich and Sokal,[4] which was used to derive (1.4)

$$M \leqslant \beta h\,\partial M/\partial(\beta h) + \beta\,|J|\,M^2(2M + M^2/h) \\ + \text{ a lower order term} \tag{1.14}$$

Such a bound follows from (1.9)–(1.10) by using first (1.12) and then (1.13) to bound the $\partial M/\partial\beta$ term. This is far from a coincidence, since, as was explained in Ref. 2, one of the two precursors of (1.11) was a percolation analogue of (1.14). The other ODI is

$$\partial M^2/\partial\beta \geqslant \text{const} \qquad (\text{for } \beta > \beta_c) \tag{1.15}$$

which generalizes a result of Aizenman.[1] A quite general percolation version of (1.15) (with a suitably reduced power of M) was obtained by Chayes and Chayes.[5] In fact, that result was the other predecessor of (1.11) alluded to above. Finally, it may be mentioned (see also Section 5) that on a technical level the new relations (1.9) and (1.10) may be regarded as special cases of general inequalities contained in the work of Aizenman and Graham.[3] However, the particular physical relations implied by those inequalities, which are discussed here, were not recognized in that work (which was mainly concerned with the renormalized coupling constant in four dimensions).

The present derivation of the inequalities (1.9) and (1.10) is based on the random current and random walk techniques of Refs. 19 and 22. In the next section we recapitulate these representations, and then apply them in Section 3 to prove the Ising case of Theorem 2. In Section 4, the argument is extended to prove the more general inequality. The consequences of Theorem 2, which include Theorem 1, are discussed in Section 5.

2. THE RANDOM CURRENT AND RANDOM WALK REPRESENTATIONS

In this section we restrict our attention to ferromagnetic models of Ising spins, and briefly review those aspects of the random current and random walk representations that are relevant for the proof of (1.9). (Other spin distributions are discussed in Section 4.)

In the absence of a magnetic field the correlation functions of Ising models can be expanded[19,22] in terms of integer-valued functions $\mathbf{n} = (n_{\{x,y\}})$ of the bonds of the lattice. Each function can be interpreted as the set of *flux numbers*, which define a "current configuration." To accommodate a (nonnegative) magnetic field in this picture, it is convenient to enhance the lattice by introducing an extra site g (playing here the role of the location of Griffiths' "ghost spin"[23]), which is linked to each lattice site x by a bond $\{x, g\}$—for which we take the coupling strength to be the external field, $J_{\{x,g\}} = h$. Since the ghost site plays a special role, we shall refer to it by the symbol g, and not the symbols used for generic lattice sites (e.g., x), unless specifically stated otherwise. Likewise, we shall distinguish between the lattice bonds $\{x, y\}$ and the "h-bonds" $\{x, g\}$.

Remark. In a previous paper[16] the formalism was presented using a layer of ghost spins rather than a single extra site. However, an inspection of the proofs shows that the arguments work in exactly the same way if all the ghost sites are merged. We adopt this simpler picture here, which is closer to the one introduced by Griffiths.

The starting point is the following representation for the partition function, obtained by factoring the Boltzmann weight into terms associated with bonds and then expanding each such factor in powers of βJ_b. After averaging over the spins one gets

$$Z \equiv 2^{-|A|} \sum_{\sigma_x = \pm 1} e^{-\beta H} = \sum_{\partial \mathbf{n} = \varnothing} w(\mathbf{n}) \tag{2.1}$$

where $\mathbf{n}$ assigns flux numbers to both lattice bonds and h-bonds b, and the weights are

$$w(\mathbf{n}) = \prod_b (\beta J_b)^{n_b}/n_b! \tag{2.2}$$

with

$$J_b = \begin{cases} J_{y-x} & \text{if } b = \{x, y\} \\ h & \text{if } b = \{x, g\} \end{cases}$$

If translation invariance is not assumed, then J_{y-x} and h should be replaced by $J_{x,y}$ and h_x, respectively. The constraint seen in the sum in (2.1) on the sets of "sources of $\mathbf{n}$," defined as

$$\partial \mathbf{n} = \left\{ x \in (-L, L]^d \cup \{g\} \,\middle|\, \sum_{b \ni x} n_b \text{ is odd} \right\} \tag{2.3}$$

For the correlation functions, which are the thermal expectation values of variables of the form $\sigma_A = \prod_{x \in A} \sigma_x$, we have

$$\langle \sigma_A \rangle = \sum_{\partial \mathbf{n} = A} w(\mathbf{n})/Z \qquad \text{if } |A| \text{ is even} \tag{2.4a}$$

and

$$\langle \sigma_A \rangle = \sum_{\partial \mathbf{n} = A \cup \{g\}} w(\mathbf{n})/Z \qquad \text{if } |A| \text{ is odd} \tag{2.4b}$$

The latter case includes of course the magnetization $\langle \sigma_x \rangle$.

Utilizing relations like (2.4), one can associate expectations of observables with probabilities of certain geometric events in a suitable system of random currents, which are just the graphs associated with expansions like (2.1) taken with their relevant weights. Of particular significance are the probabilities of the existence or absence of various connections. We adopt here the natural notion of connection: two sites x and y (here y may be the ghost site) are said to be *connected* in the current configuration $\mathbf{n}$ if there exists a path of bonds with $n_b \neq 0$ joining x with y. In such a case we shall write $\mathbf{n}: x \leftrightarrow y$. Having defined when two sites are connected, we can introduce the notion of a cluster. The *(bond) cluster* of a site x (x may be the ghost site) in a current configuration $\mathbf{n}$ is the set of bonds with at least one site connected to x:

$$C_{\mathbf{n}}(x) = \{\{y, z\} \mid \mathbf{n}: y \leftrightarrow x \text{ or } \mathbf{n}: z \leftrightarrow x\} \cup \{\{y, g\} \mid \mathbf{n}: y \leftrightarrow x\}$$

In particular, $C_{\mathbf{n}}(g)$ contains at least all of the "h-bonds." Those bonds for which only one endpoint is connected to x (i.e., "dangling bonds") will be referred to as the boundary of $C_{\mathbf{n}}(x)$.

One of the reasons for the usefulness of this random current representation is the existence of a combinatorial *identity* which allows us to switch sources in a *duplicated* system of currents.[19,22] The particular case of this switching lemma that we shall employ is the following.

Lemma 2.1. Let A be a set of sites of the enhanced lattice, $\{x, y\}$ a pair of sites (one of which may be g), and f a function defined on current configurations. Then

$$\sum_{\partial \mathbf{n}_1 = A, \partial \mathbf{n}_2 = \{x,y\}} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2)$$
$$= \sum_{\partial \mathbf{n}_1 = A \Delta \{x,y\}, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2)\, I[\mathbf{n}_1 + \mathbf{n}_2 : x \leftrightarrow y] \tag{2.5}$$

Here Δ stands for the symmetric difference $[A \Delta B = (A \cup B) \backslash (A \cap B)]$, $\mathbf{n}_1 + \mathbf{n}_2$ is defined by taking the "bondwise" sum of the two currents $[(\mathbf{n}_1 + \mathbf{n}_2)_b = n_{1b} + n_{2b}]$, and $I[Q]$ is the indicator function, which takes the value 1 if the condition Q is satisfied and 0 otherwise.

This lemma is especially useful in dealing with truncated correlation functions, where often it precisely accounts for delicate cancellations among contributions of different sign. Two examples we use later are the following[16]:

$$\langle \sigma_0 ; \sigma_x \rangle = \sum_{\partial \mathbf{n}_1 = \{0\} \Delta \{x\}, \partial \mathbf{n}_2 = \varnothing} Z^{-2}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2) \times I[\mathbf{n}_1 + \mathbf{n}_2 : 0 \not\leftrightarrow g] \tag{2.6}$$

and

$$\langle \sigma_0 ; \sigma_u \sigma_v \rangle = \sum_{\partial \mathbf{n}_1 = \{0,g\} \Delta \{u\} \Delta \{v\}, \partial \mathbf{n}_2 = \varnothing} Z^{-2}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2) \times I[\mathbf{n}_1 + \mathbf{n}_2 : 0 \not\leftrightarrow g] \tag{2.7}$$

The truncated correlation functions appearing in (2.6) and (2.7) will enter our analysis through the differentiation formulas

$$\chi = \partial M/\partial(\beta h) = \sum_x \langle \sigma_0; \sigma_x \rangle \tag{2.8}$$

and

$$\partial M/\partial\beta = \tfrac{1}{2} \sum_{x,y} J_{x-y} \langle \sigma_0; \sigma_x \sigma_y \rangle \tag{2.9}$$

Formulas (2.6) and (2.7) can be written in a somewhat more detailed manner by "conditioning on clusters." For instance, in (2.7) the source constraints together with the indicator function imply that either u is the endpoint of a bond in $C_{\mathbf{n}_1+\mathbf{n}_2}(0)$ and v is not, or vice versa. This fact can be represented geometrically as in Fig. 1 (cf. Fig. 1 of Ref. 2). When conditioned on the cluster $C_{\mathbf{n}_1+\mathbf{n}_2}(0)$, the sums over the current configurations inside and outside the cluster are conditionally independent. The exterior sum gives the expectation of σ_u (or σ_v) in the system that has been deprived of the bonds in $C_{\mathbf{n}_1+\mathbf{n}_2}(0)$, multiplied by the partition function in this weakened system. Using this partition function to "uncondition" the sum, we obtain

$$\langle \sigma_0; \sigma_u \sigma_v \rangle = \sum_{\partial \mathbf{n}_1 = \{0\} \Delta \{u\}, \partial \mathbf{n}_2 = \varnothing} Z^{-2} w(\mathbf{n}_1) \, w(\mathbf{n}_2) \langle \sigma_v \rangle_{[C(0)]^c} + \text{a } u \Leftrightarrow v \text{ permutation} \tag{2.10}$$

We have written $C(0)$ as an abbreviation for $C_{\mathbf{n}_1+\mathbf{n}_2}(0)$ and we follow the convention that if A is a set of bonds, then the subscript A, as in Z_A and $\langle \cdot \rangle_A$, indicates that the coupling constants are set to zero for all bonds not belonging to A. A similar argument yields

$$\langle \sigma_0; \sigma_x \rangle = \sum_{\partial \mathbf{n}_1 = \varnothing, \partial \mathbf{n}_2 = \varnothing} Z^{-2} w(\mathbf{n}_1) \, w(\mathbf{n}_2) \langle \sigma_0 \sigma_x \rangle_{[C(g)]^c} \times I[\mathbf{n}_1 + \mathbf{n}_2 \colon 0 \nleftrightarrow g] \tag{2.11}$$

(The indicator function is superfluous except in the case $x = 0$.)

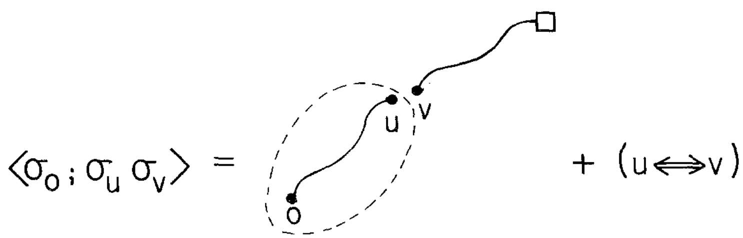

Fig. 1. Diagrammatic representation of (2.7). The dashed line stands for the boundary of $C_{\mathbf{n}_1+\mathbf{n}_2}(0)$, each solid line indicates a "backbone" (defined below), and the square symbolizes an h-connection to the ghost site g.

For some purposes it is convenient to resum the random current expansion into a more compact-looking random walk expansion. This is accomplished by observing that a current configuration $\mathbf{n}$ with $\partial\mathbf{n} = B$ (B may include the ghost site g) must exhibit a family of walks on bonds having *odd* flux that connect the elements of B in pairs. By specifying an appropriate set of rules, these walks can be assigned uniquely so as to create a well-defined map associating a sequence of walks $\Omega_B(\mathbf{n})$ to each current configuration $\mathbf{n}$ with $\partial\mathbf{n} = B$. We refer to $\Omega_B(\mathbf{n})$ as the *B-backbone* of $\mathbf{n}$. A partial resumation of (2.4) yields

$$\langle \sigma_A \rangle = \sum_{\partial\omega = A} \rho(\omega) \qquad \text{if} \quad |A| \text{ is even} \tag{2.12a}$$

$$\langle \sigma_A \rangle = \sum_{\partial\omega = A \cup \{g\}} \rho(\omega) \qquad \text{if} \quad |A| \text{ is odd} \tag{2.12b}$$

where

$$\rho(\omega) = \sum_{\partial\mathbf{n} = \partial\omega} Z^{-1} w(\mathbf{n})\, I[\Omega_{\partial\omega}(\mathbf{n}) = \omega] \tag{2.13}$$

and $\partial\omega$ is defined by the natural analogue of (2.3)—it is the set of points in $(-L, L]^d \cup \{g\}$ visited by an odd number of steps of ω.

Before we describe the algorithmic construction used here to define the backbone, let us present the main properties of the weights $\rho(\omega)$ that the construction yields for the above expansion.

There will be a rule associating to each sequence of walks ω a collection of bonds $\tilde{\omega}$, referred to as the set of bonds canceled by ω, which in general will contain the set of bonds traversed by the steps in ω and some of their neighbors. The sets $\tilde{\omega}$ satisfy:

(i) If for some configuration $\mathbf{n}$, with $\partial\mathbf{n} = B$, the backbone is ω, then $\Omega_B(\mathbf{n}') = \omega$ for every other configuration $\mathbf{n}'$ that coincides with $\mathbf{n}$ on $\tilde{\omega}$, and also has $\partial\mathbf{n}' = B$.

The weights $\rho(\omega)$ that enter in the expansion (2.12) vanish for certain "inconsistent" sequences of walks. The consistency criterion imposes constraints both within the individual walks in ω and among them, and furthermore on the order in which they appear in ω. Two key properties of the weights are as follows.

(ii) For every decomposition of a consistent sequence of walks ω into two sequences $\omega = \omega_1 \circ \omega_2$ (possibly through the "cutting" of one of the walks of ω into two parts)

$$\rho(\omega_1 \circ \omega_2) = \rho(\omega_1)\, \rho_{\omega_1^c}(\omega_2) \tag{2.14}$$

where $\rho_{A^c}(\omega)$ is the weight for the system in which all the coupling constants of the bonds in A are set to zero.

(iii) For any set of bonds A, if ω is a sequence of walks none of which passes through a bond in A (i.e., $\omega \cap A = \varnothing$), then

$$\rho(\omega) \leqslant \rho_{A^c}(\omega) \tag{2.15}$$

The last inequality can be interpreted as saying that the weights $\rho(\omega)$ *decrease* when the Hamiltonian is enhanced by new bond interactions, which (by the Griffiths inequalities) is the opposite of what happens to the total sums in (2.12). There is no contradiction here, since new interaction terms increase the collection of walks, by adding connections in the relevant graph.

It should be noted that there are at present various random walk expansions, which share a number of useful properties.[19,22,24] The backbone expansion used here is conveniently related to the more explicit random current representation, allowing us to get further advantage by mixing the two. Such mixed arguments appeared already in Refs. 3 and 19; however we shall use here a slightly modified construction (and hence also definition) of the backbone, following Ref. 22. Let us now give the algorithm for the construction of the backbone (which is nonunique because of the choice of an order).

We start by ordering the (countable) lattice so that for each subset there is an "earliest" site. (Although the "lexicographic order" does not have this property, it would have sufficed for finite-range models.) The order is extended to the enhanced lattice by declaring the ghost site to be the earliest of all. We shall now construct walks consisting of *steps*, where a *step* is an ordered pair of sites (recall that bonds are unordered pairs), of which the first must belong to the lattice and the second may be either a lattice site or the ghost site. With each step (x, v) we associate a set of *canceled* bonds, which consists of all bonds $\{x, y\}$ with sites y in the enhanced lattice that are "earlier" than v.

For a current configuration $\mathbf{n}$ with $\partial\mathbf{n} = B$, the backbone $\Omega_B(\mathbf{n})$ consists of a number of walks constructed as follows.

A1. The first walk, ω_1, starts from the earliest of the lattice sites in $B \setminus \{g\}$, which we refer to as u_1, and stops upon reaching either another site in B or the ghost site (even if g is not in B). Its first step is (u_1, u_2), where u_2 is the earliest site v for which the flux number $n_{\{u_1, v\}}$ is odd.

A2. Until the walk stops, each step (u_i, u_{i+1}) is the earliest of those steps (u_i, v) emerging from u_i that does not use bonds canceled by any of the previous steps and for which $n_{\{u_i, v\}}$ is odd.

A3. When ω_1 stops, we pick the first of the lattice sites of B not visited by ω_1, and repeat the process, excluding, in the choice of the steps, all the bonds canceled by ω_1. This defines a second walk ω_2.

A4. Continue until B is exhausted. At the end we obtain walks ω_1, $\omega_2,..., \omega_s$ $(s \leqslant |B|)$ traversed in that order. We call this set of walks the *B-backbone* and denote it by $\Omega_B(\mathbf{n}) = \omega_1 \circ \omega_2 \circ \cdots \circ \omega_s$.

It is easy to see that the backbone will always satisfy the following consistency conditions:

B1. The starting points of the walks ω_i are increasing in our order.

B2. No step uses a bond canceled by a previous step.

We define $\tilde{\omega}$ as the set of bonds canceled by the steps in ω. It should be noted that among the configurations $\mathbf{n}$ having $\partial\mathbf{n} = B$, those whose backbone is ω are characterized by the following requirements, which refer only to the restriction of $\mathbf{n}$ to $\tilde{\omega}$:

C1. $n_{\{x,y\}}$ is odd for all $(x, y) \in \omega$.

C2. $n_{\{x,y\}}$ is even for all $(x, y) \in \tilde{\omega} \backslash \omega$ [had such an $n_{\{x,y\}}$ been odd, then the backbone would have gone through (x, y) by A2].

The key properties (2.14) and (2.15) are derived by the arguments of Ref. 19; for the algorithm used here the proofs can be found in Ref. 16.

To make the pairing of sources associated with the backbone more explicit, we shall use the notation ω_i: $x \to y$ to mean that ω_i is a path that starts at x, ends at y, and visits y only *once*. For example,

$$\langle \sigma_0 \rangle = \sum_{\omega:0 \to g} \rho(\omega) \tag{2.16}$$

but

$$\langle \sigma_x \sigma_y \rangle = \sum_{\omega:x \to y} \rho(\omega) \tag{2.17}$$

only in the *absence* of magnetic field. If there is a magnetic field, we must add a term $\sum_{\omega_1:x \to g, \omega_2:y \to g} \rho(\omega_1 \circ \omega_2)$ to the right-hand side of (2.17). The partial contribution found on the right-hand side of (2.17) does, however, define a very useful "kernel" K, of the type first introduced by Fröhlich and Sokal[4] and used also in Ref. 16.

3. PROOF OF THE MAIN INEQUALITY

In this section we derive the Ising spin case (1.9) of our new inequality, which is repeated below.

Theorem 2(i). For a translation-invariant ferromagnetic Ising spin model in $(-L, L]^d$, the magnetization obeys the bound

$$M \leqslant \tanh(\beta h)\chi + M^3 + M^2\beta\, \partial M/\partial\beta \tag{3.1}$$

Remark. For non-translation-invariant systems, we obtain the bound

$$M_x \leqslant \sum_v \tanh(\beta h_v)\, \partial M_x/\partial(\beta h_v) + M_x^3 + \sum_{u,v} \tanh(\beta J_{u,v}) M_v^2\, \partial M_x/\partial(\beta J_{u,v}) \tag{3.2}$$

Before turning to the proof, let us describe the general idea behind it. The reader is advised, however, to regard the description given next only as an impressionistic picture.

To set the intuition on the right track, let us consider the representation

$$\langle\sigma_0\rangle = \sum_{\partial \mathbf{n}_1 = \{0, g\}} w(\mathbf{n}_1)/Z \tag{3.3}$$

with the weights given by (2.2). Note that in the high-temperature regime, $\beta \ll \beta_c$, the fluxes of the contributing configurations will seldomly take values other than 0 or 1. Likewise, the overlap of two independent current configurations, as in (2.5), will be very sparse unless the two configurations are constrained to have nearby sources.

Each current appearing in the sum in (3.3) contains a current line from 0 to the ghost site g. Note that, by parity constraints, the flux from the cluster $C(0)$ to g is either one or at least three. It is instructive now to divide the flux configurations $\{n_b\}$ appearing in (3.3) into the following three classes.

1. The flux into g from the connected cluster of 0 [as measured by the sum of n_b over the h-bonds from $C(0)$] is exactly one.

2. The origin is "triply" connected to g, in the sense that there are three flux lines each leading from 0 to g for which the sum of the fluxes is dominated by the given flux configuration $\{n_b\}$.

3. There is a "pivotal bond," $b = \{u, v\}$, such that $n_b = 1$, the origin 0 is connected to u, v is "triply" connected to g, and any connection from u to g passes through b.

The above decomposition is quite similar to the one used in the analysis of the order parameter in percolation models,[2] except for the "triple" connection appearing here in place of the "double" connection, which is all that one may expect in the absence of any source constraints.

When the overlap effects are ignorable, the leading orders for each of the three partial sums in the above decomposition of $M = \langle\sigma_0\rangle$ are (1) $\beta h\, \partial M/\partial(\beta h)$, (2) M^3, and (3) $M^2\, \partial M/\partial\beta$, which add up to (3.1). Of course, this observation may be of no value, since in this regime the first term greatly overwhelms the other two (which is certainly not the case in the vicinity of the critical point).

While the neglect of the various overlap effects is totally unjustified in the nonperturbative regime in which we are interested, we are able to present a valid "dressed" version of the above argument, which, however, requires a somewhat more complicated setup.

Proof. Our argument uses the analysis of various terms in a *tripled* system of currents. The starting point is the following expression, which is obtained by adding two dummy summations to (3.3):

$$\langle\sigma_0\rangle = \sum_{\partial\mathbf{n}_1 = \{0,g\}, \partial\mathbf{n}_2 = \varnothing, \partial\mathbf{n}_3 = \varnothing} Z^{-3}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3) \tag{3.4}$$

We shall be interested in the lattice points of the $\{0, g\}$-backbone of $\mathbf{n}_1$, $\Omega_{\{0,g\}}(\mathbf{n}_1)$, which are connected to g via $\mathbf{n}_2 + \mathbf{n}_3$. Classifying the current configurations by the first such site (where "first" means with respect to the natural order induced by the direction of "travel" along the backbone), we distinguish three cases corresponding to the three classes of configurations given in the heuristic discussion above.

1. There is no such site.
2. The first such site is 0.
3. The first such site is some $v \neq 0$.

In the third case, the step (u, v) at which v is first reached divides the backbone into two parts, the first of which links 0 to u without being connected to g by the $\mathbf{n}_2 + \mathbf{n}_3$ current configuration.

On the basis of the above decomposition, we write for $M = \langle\sigma_0\rangle$

$$M = T + R_0 + \sum_u \sum_{v\neq 0} R_{uv} \tag{3.5}$$

with

$$T = \sum_{\partial\mathbf{n}_1 = \{0,g\}, \partial\mathbf{n}_2 = \varnothing, \partial\mathbf{n}_3 = \varnothing} Z^{-3}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times\ I[\text{none of the lattice bonds of } \Omega_{\{0,g\}}(\mathbf{n}_1) \text{ is in } C_{\mathbf{n}_2+\mathbf{n}_3}(g)] \tag{3.6}$$

$$R_0 = \sum_{\partial\mathbf{n}_1 = \{0,g\}, \partial\mathbf{n}_2 = \varnothing, \partial\mathbf{n}_3 = \varnothing} Z^{-3}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times\ I[\mathbf{n}_2 + \mathbf{n}_3 \colon 0 \leftrightarrow g] \tag{3.7}$$

and

$$R_{uv} = \sum_{\partial \mathbf{n}_1 = \{0,g\}, \partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-3} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times\, I[\mathbf{n}_2 + \mathbf{n}_3 \colon v \leftrightarrow g]\; I[(u, v) \in \Omega_{\{0,g\}}(\mathbf{n}_1)]$$
$$\times\, I[\text{none of the bonds of } \Omega_{\{0,g\}}(\mathbf{n}_1) \text{ prior to } \{u, v\} \text{ is in } C_{\mathbf{n}_2 + \mathbf{n}_3}(g)] \tag{3.8}$$

Due to the source constraints, every site connected to g via $\mathbf{n}_2 + \mathbf{n}_3$ must in fact be "doubly" connected. Hence the decomposition (3.5) may be depicted as in Fig. 2.

Our purpose now is to bound the above terms by "physical quantities." This will be done by combining random walk techniques, applied to the backbone of $\mathbf{n}_1$, with random current techniques, applied to $\mathbf{n}_2$ and $\mathbf{n}_3$.

First, let us show that

$$T \leqslant \tanh(\beta h)\, \partial M / \partial(\beta h) \tag{3.9}$$

To prove (3.9), we classify the configurations of the right-hand side of (3.6) by the last step (x, g) of the $\mathbf{n}_1$ backbone and observe that the flux $n_{1\{x,g\}}$ must be *odd*, and that this parity constraint is the only effect of conditioning on the values of all the other flux numbers. For each site x we estimate the corresponding contribution to T by considering the set of currents obtained by flipping this parity constraint on $n_{1,\{x,g\}}$. The sum with the *odd* constraint clearly equals the sum with the reversed parity multiplied by the factor $\tanh(\beta h)$. For the new sum, we note that its configurations $\mathbf{n}_1$ have the sources $\{0, x\}$, and that they satisfy the condition that the $\mathbf{n}_1$ backbone reaches x before hitting the $\mathbf{n}_2 + \mathbf{n}_3$ cluster of the ghost site g, $C_{\mathbf{n}_2 + \mathbf{n}_3}(g)$.

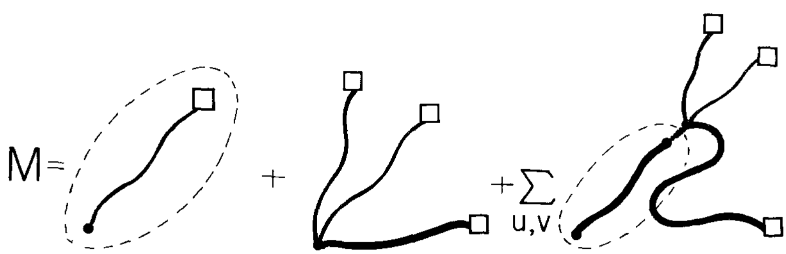

Fig. 2. Diagrams representing the different contributions to M in (3.5). The heavy line is the backbone $\Omega_{\{0,g\}}(\mathbf{n}_1)$, other solid lines represent connections due to $\mathbf{n}_2 + \mathbf{n}_3$ and the dashed surface represents the boundary of $C_{\mathbf{n}_2 + \mathbf{n}_3}(g)$, which is not necessarily connected.

Hence,

$$T = \tanh(\beta h) \sum_x \sum_{\partial \mathbf{n}_1 = \{0\} \Delta \{x\}, \partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-3} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times I[n_{1\{0,x\}} \text{ is even}]\, I[\Omega_{\{0,x\}}(\mathbf{n}_1) \cap C_{\mathbf{n}_2+\mathbf{n}_3}(g) = \varnothing]$$
$$\times I[\mathbf{n}_2 + \mathbf{n}_3 : 0 \not\leftrightarrow g] \tag{3.10}$$

(The last indicator function is included to cover the special case $x = 0$.) Bounding the first indicator function by 1 and using the random walk representation for $\mathbf{n}_1$, one obtains

$$T \leqslant \tanh(\beta h) \sum_x \sum_{\partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-2} w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times I[\mathbf{n}_2 + \mathbf{n}_3 : 0 \not\leftrightarrow g]$$
$$\times \sum_{\omega : 0 \to x} \rho(\omega)\, I[\omega \cap C_{\mathbf{n}_2+\mathbf{n}_3}(g) = \varnothing]$$
$$\leqslant \tanh(\beta h) \sum_x \sum_{\partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-2} w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times I[\mathbf{n}_2 + \mathbf{n}_3 : 0 \not\leftrightarrow g]$$
$$\times \sum_{\omega : 0 \to x} \rho_{[C(g)]^c}(\omega)\, I[\omega \cap C_{\mathbf{n}_2+\mathbf{n}_3}(g) = \varnothing] \tag{3.11}$$

where the last inequality is by the monotonicity property (2.15). Finally, we use (2.17) [the system deprived of the bonds of $C_{\mathbf{n}_2+\mathbf{n}_3}(g)$ has no magnetic field] and (2.11) to get

$$T \leqslant \tanh(\beta h) \sum_x \sum_{\partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-2} w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times \langle \sigma_0 \sigma_x \rangle_{[C(g)]^c}\, I[\mathbf{n}_2 + \mathbf{n}_3 : 0 \not\leftrightarrow g]$$
$$= \tanh(\beta h) \sum_x \langle \sigma_0 ; \sigma_x \rangle$$

which, by (2.8), proves (3.9).

For the next term in (3.5) we get, by an easy application of the switching lemma (2.5), the formula

$$R_0 = \sum_{\partial \mathbf{n}_1 = \{0, g\}, \partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-3} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$
$$\times I[\mathbf{n}_2 + \mathbf{n}_3 : 0 \leftrightarrow g]$$
$$= \langle \sigma_0 \rangle^3 \tag{3.12}$$

where one factor $\langle\sigma_0\rangle$ comes from the sum over $\mathbf{n}_1$ (which is independent of $\mathbf{n}_2$ and $\mathbf{n}_3$) and the other two reflect the fact that the sourceless character of $\mathbf{n}_2+\mathbf{n}_3$ forces 0 to be "doubly" connected to g whenever the two are connected.

Let us now turn our attention to the terms R_{uv} appearing in the last sum in (3.5), for which we combine the methods used to treat the preceding two terms. In each configuration of currents contributing in (3.8) the flux $n_{1\{u,v\}}$ is odd. Changing its parity, as we did with $n_{1\{x,g\}}$ in the above analysis of T, we obtain a sum that is multiplied by the factor $\tanh(\beta J_{v-u})$. The resulting configurations $\mathbf{n}_1$ have sources $\{0\}\,\Delta\{u,v,g\}$ and their $\{0\}\,\Delta\{u,v,g\}$-backbones are of the form $\omega_1\circ\omega_2$ with $\omega_1\colon 0\to u$ completely contained in $C_{\mathbf{n}_2+\mathbf{n}_3}(g)^c$ and $\omega_2\colon v\to g$. Therefore,

$$\begin{aligned} R_{uv}\leqslant \tanh(\beta J_{v-u}) &\sum_{\partial\mathbf{n}_2=\varnothing,\partial\mathbf{n}_3=\varnothing} Z^{-2}\,w(\mathbf{n}_2)\,w(\mathbf{n}_3)\\ &\times I[\mathbf{n}_2+\mathbf{n}_3\colon 0\nleftrightarrow g]\,I[\mathbf{n}_2+\mathbf{n}_3\colon v\leftrightarrow g]\\ &\times\left\{\sum_{\omega_1:0\to u,\omega_2:v\to g}\rho(\omega_1\circ\omega_2)\,I[\omega_1\cap C_{\mathbf{n}_2+\mathbf{n}_3}(g)=\varnothing]\right\} \end{aligned}\tag{3.13}$$

The quantity inside the braces in (3.13) may be bounded as follows:

$$\begin{aligned} \{\cdots\} &= \sum_{\omega_1:0\to u}\rho(\omega_1)\,I[\omega_1\cap C_{\mathbf{n}_2+\mathbf{n}_3}(g)=\varnothing]\sum_{\omega_2:v\to g}\rho_{\tilde\omega_1^c}(\omega_2)\\ &\leqslant \sum_{\omega_1:0\to u}\rho_{[C(g)]^c}(\omega_1)\langle\sigma_v\rangle_{\tilde\omega_1^c}\\ &\leqslant \langle\sigma_v\rangle\langle\sigma_0\sigma_u\rangle_{[C(g)]^c} \end{aligned}$$

where we applied first (2.14), then (2.15), and finally (2.17) and Griffiths II. [Again, $C(g)=C_{\mathbf{n}_2+\mathbf{n}_3}(g)$.] Substituting this bound into (3.13), we obtain

$$\begin{aligned} R_{uv}\leqslant \tanh(\beta J_{v-u})\langle\sigma_v\rangle &\sum_{\partial\mathbf{n}_2=\varnothing,\partial\mathbf{n}_3=\varnothing} Z^{-2}\,w(\mathbf{n}_2)\,w(\mathbf{n}_3)\\ &\times\langle\sigma_0\sigma_u\rangle_{[C(g)]^c} I[\mathbf{n}_2+\mathbf{n}_3\colon 0\nleftrightarrow g]\,I[\mathbf{n}_2+\mathbf{n}_3\colon v\leftrightarrow g]\\ =\tanh(\beta J_{v-u})\langle\sigma_v\rangle &\sum_{\partial\mathbf{n}_2=\{v,g\},\partial\mathbf{n}_3=\{v,g\}} Z^{-2}w(\mathbf{n}_2)\,w(\mathbf{n}_3)\\ &\times\langle\sigma_0\sigma_u\rangle_{[C(g)]^c} I[\mathbf{n}_2+\mathbf{n}_3\colon 0\nleftrightarrow g] \end{aligned}\tag{3.14}$$

the last step being an application of the switching lemma (2.5).

The sum of the right-hand side of (3.14) can be viewed as an expression obtained through the conditioning on the cluster of the ghost site $C_{\mathbf{n}_2+\mathbf{n}_3}(g)$ of the following quantity:

$$\sum_{\partial\mathbf{n}_2=\{0\}\Delta\{u,v,g\},\partial\mathbf{n}_3=\{v,g\}} Z^{-2}\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)\, I[\mathbf{n}_2+\mathbf{n}_3: 0 \not\leftrightarrow g]$$

$$= \sum_{\partial\mathbf{n}_2=\{0\}\Delta\{u\},\partial\mathbf{n}_3=\varnothing} Z^{-2}\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)(\langle\sigma_v\rangle_{[C(0)]^c})^2 \qquad (3.15)$$

where on the right-hand side of (3.15) we have "turned the conditioning around," onto $C(0)=C_{\mathbf{n}_2+\mathbf{n}_3}(0)$.

We now insert (3.15) into (3.14) and use Griffiths II to bound one of the factors $\langle\sigma_v\rangle_{[C(0)]^c}$ by the full magnetization $\langle\sigma_v\rangle = M$, getting

$$R_{uv} \leqslant \tanh(\beta J_{v-u})\, M^2 \sum_{\partial\mathbf{n}_2=\{0\}\Delta\{u\},\partial\mathbf{n}_3=\varnothing} Z^{-2}\, w(\mathbf{n}_2)\, w(\mathbf{n}_3)$$

$$\times\ \langle\sigma_v\rangle_{[C(g)]^c} \qquad (3.16)$$

The sum on the right-hand side of (3.17) is one of the two terms, which differ only by a $u \Leftrightarrow v$ permutation, appearing in the expansion (2.10) of $\langle\sigma_0;\sigma_u\sigma_v\rangle$. Therefore, we have obtained

$$\sum_{u,v} R_{uv} \leqslant \tfrac{1}{2} M^2 \sum_{u,v} \tanh(\beta J_{v-u})\langle\sigma_0;\sigma_u\sigma_v\rangle \qquad (3.17)$$

Combining now the bounds (3.9), (3.12), and (3.17) that we have for the three terms in (3.5), and replacing $\tanh(\beta J_{v-u})$ with βJ_{v-u}, we obtain (3.1). ∎

4. EXTENSION TO SPINS IN THE GRIFFITHS–SIMON CLASS

We shall now derive the differential inequality (1.10) for the more general case of spins whose *a priori* measure ρ_0 belongs to the Griffiths–Simon class. Generally speaking, spins in that class are representable as "block variables" of ferromagnetic ensembles of Ising spins. To distinguish between the two levels of description, we shall denote here the original model's spin variables as φ_x. We make the standing assumption that $\int \exp(a\varphi^2)\,\rho_0(d\varphi) < \infty$ for all $a<\infty$. It should be noted that the class of measures we are discussing now includes continuous as well as discrete spins, e.g., the "φ^4 variables" with $\rho_0(\varphi) = \exp(-\lambda\varphi^4 + b\varphi^2)$.[7]

The basic construction consists in writing the spin φ as a weighted sum of "microscopic" Ising spins σ_α

$$\varphi = \sum_{1\leqslant\alpha\leqslant N} Q_\alpha^{(N)}\sigma_\alpha$$

where $Q_\alpha^{(N)}$ are some positive coefficients and the spins σ_α are ferromagnetically coupled via a Hamiltonian of the form

$$H_1^{(N)} = -\frac{1}{2} \sum_{\alpha,\delta} I_{\alpha,\delta}^{(N)} \sigma_\alpha \sigma_\delta \tag{4.1}$$

The resulting distribution for the variable φ is in the GS class, which is further augmented by the collection of all distributional limits (as $N \to \infty$) of such measures.

When dealing with a lattice system it is convenient to use the above representation at each site. Thus, we add a second label α $(=1,..., N)$ in addition to the lattice site, and consider the array of "microscopic" sites $\{(x, \alpha): x \in (-L, L]^d; \alpha = 1,..., N\}$ with the spins $\sigma_{(x,\alpha)}$. A lattice site x is represented now by the "block" $B_x = \{(x, \alpha): \alpha = 1,..., N\}$, and the spin variables φ_x are written as

$$\varphi_x = \sum_{1 \leqslant \alpha \leqslant N} Q_\alpha^{(N)} \sigma_{(x,\alpha)} \tag{4.2}$$

Therefore, systems with GS spins and a macroscopic two-body interaction

$$H_2 = -\frac{1}{2} \sum_{x,y} J_{x-y} \varphi_x \varphi_y - h \sum_x \varphi_x \tag{4.3}$$

can be approximated by Ising models with total Hamiltonians of the form

$$\begin{aligned} H^{(N)} &= H_1^{(N)} + H_2 \\ &= -\frac{1}{2} \sum_{x,y,\alpha,\delta} J_{\{(x,\alpha),(y,\delta)\}} \sigma_{(x,\alpha)} \sigma_{(y,\delta)} - \sum_\alpha h_\alpha \sigma_{(x,\alpha)} \end{aligned} \tag{4.4}$$

where

$$J_{\{(x,\alpha),(y,\delta)\}} = \begin{cases} J_{x,y} Q_\alpha^{(N)} Q_\delta^{(N)} & \text{if } \ y \neq x \\ I_{x,\alpha,\delta}^{(N)} & \text{if } \ y = x \end{cases} \tag{4.5a}$$

and

$$h_\alpha = h Q_\alpha^{(N)} \tag{4.5b}$$

The above representation is especially useful if one succeeds in deriving, from the random current representation for the underlying spins σ, relations that can be expressed in terms of correlation functions of the "block spins" φ. Within the framework of the random current representation, the crucial step in obtaining such expressions is the replacement of *suitably chosen* connection constraints involving microscopic *sites* by constraints referring to macroscopic *blocks*.

Following these guidelines, we shall prove in this section the extension of our main inequality to GS spins with the ferromagnetic Hamiltonian (4.3). The approximating Ising model with Hamiltonian (4.4) has the following random current representation for its order parameter:

$$M = \sum_{\alpha} Q_{\alpha}^{(N)} \sum_{\partial \mathbf{n}_1 = \{(0,\alpha), g\}} w(\mathbf{n}_1)/Z \tag{4.6}$$

As in Section 3, it is useful to consider an expression for M where two dummy summations are included in the expansion (4.6), and then to apply a decomposition similar to the one used there. The principal difference is that we shall replace sites by blocks in the various connection constraints, which will require the following block versions of the switching lemma (2.5).

Given a collection of sites B and a current configuration $\mathbf{n}$, we shall say that $\mathbf{n}$ connects B with another site z, and denote this condition by $\mathbf{n}$: $B \leftrightarrow z$, if there exists an $x \in B$ such that $\mathbf{n}$: $x \leftrightarrow z$.

Lemma 4.1. Let B be a collection of sites and z a site outside B. Then, for any function f on current configurations that is decreasing in each flux number,

$$\sum_{\partial \mathbf{n}_1 = \varnothing, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2)\, I[\mathbf{n}_1 + \mathbf{n}_2 \colon B \leftrightarrow z]$$

$$\leqslant \sum_{x \in B, y \in B^c} \beta J_{y-x} \sum_{\partial \mathbf{n}_1 = \{x,z\}, \partial \mathbf{n}_2 = \{y\} \Delta \{z\}} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2) \tag{4.7}$$

Remark. A more general version of this lemma (with $\partial \mathbf{n}_1 = \{u\} \Delta \{v\}$ and without the monotonicity requirement on f) is proven in Ref. 3, Proposition 7.1. The proof of (4.7) presented here is, however, considerably simpler than that of the general version and exhibits more clearly the essential idea.

Proof. The key observation is that as $\mathbf{n}_1 + \mathbf{n}_2$ is a sourceless current, every connection from B to z must form a "loop," i.e., the total flux through bonds linking B with B^c must be at least 2. Therefore

$$I[\mathbf{n}_1 + \mathbf{n}_2 \colon B \leftrightarrow z] \leqslant \sum_{x \in B, y \in B^c} [(n_{1\{x,y\}} + n_{2\{x,y\}})/2] \times I[\mathbf{n}_1 + \mathbf{n}_2 |_{\mathbf{B}^c} \colon y \leftrightarrow z] \tag{4.8}$$

where we have denoted by **B** the set of bonds with at least one endpoint in B. Inequality (4.8) allows us to bound the left-hand side (LHS) of (4.7):

$$\begin{aligned} \text{LHS} \leqslant \tfrac{1}{2} \sum_{x \in B, y \in B^c} \sum_{\partial \mathbf{n}_1 = \varnothing, \partial \mathbf{n}_2 = \varnothing} & \{ (n_{1\{x,y\}} w(\mathbf{n}_1)\, w(\mathbf{n}_2) \\ & \times f(\mathbf{n}_1 + \mathbf{n}_2)\, I[(\mathbf{n}_1 + \mathbf{n}_2)|_{\mathbf{B}^c}\colon y \leftrightarrow z]) \\ & + (\mathbf{n}_1 \Leftrightarrow \mathbf{n}_2 \text{ permutation}) \} \end{aligned} \tag{4.9}$$

We next note that the second term, corresponding to the $\mathbf{n}_1 \Leftrightarrow \mathbf{n}_2$ permutation, yields the same contribution as the first term and hence it can be eliminated by replacing the prefactor 1/2 with 1. Moreover, from (2.2) we see that

$$\begin{aligned} n_{1\{x,y\}} w(\mathbf{n}_1) &= [(\beta J_{y-x})^{n_{1\{x,y\}}} / (n_{1\{x,y\}} - 1)!] \\ & \quad \times \prod_{b \neq \{x,y\}} (\beta J_b)^{n_b} / n_b!] \\ &= \beta J_{y-x} w(\mathbf{n}_1 - \delta_{b,\{x,y\}}) \end{aligned} \tag{4.10}$$

Hence, replacing the current configuration $\mathbf{n}_1$ with the configuration obtained by reducing its flux across $\{x, y\}$ by 1, i.e., $\mathbf{n}_1 - \delta_{b,\{x,y\}}$, we obtain

$$\begin{aligned} \text{LHS} \leqslant \sum_{x \in B, y \in B^c} \beta J_{y-x} & \sum_{\partial \mathbf{n}_1 = \{x,y\}, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2) \\ & \times f(\mathbf{n}_1 + \mathbf{n}_2 + \delta_{b,\{x,y\}})\, I[(\mathbf{n}_1 + \mathbf{n}_2 + \delta_{b,\{x,y\}})|_{\mathbf{B}^c}\colon y \leftrightarrow z] \end{aligned} \tag{4.11}$$

We now bound each of the last two factors in the sum in (4.11). From the monotonicity assumption on f,

$$f(\mathbf{n}_1 + \mathbf{n}_2 + \delta_{b,\{x,y\}}) \leqslant f(\mathbf{n}_1 + \mathbf{n}_2)$$

and because $\{x, y\} \notin \mathbf{B}^c$,

$$\begin{aligned} I[(\mathbf{n}_1 + \mathbf{n}_2 + \delta_{b,\{x,y\}})|_{\mathbf{B}^c}\colon y \leftrightarrow z] &= I[(\mathbf{n}_1 + \mathbf{n}_2)|_{\mathbf{B}^c}\colon y \leftrightarrow z] \\ &\leqslant I[\mathbf{n}_1 + \mathbf{n}_2\colon y \leftrightarrow z] \end{aligned}$$

Combining these two observations, one obtains from (4.11)

$$\begin{aligned} \text{LHS} \leqslant \sum_{x \in B, y \in B^c} \beta J_{y-x} & \sum_{\partial \mathbf{n}_1 = \{x,y\}, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2) \\ & \times f(\mathbf{n}_1 + \mathbf{n}_2)\, I[\mathbf{n}_1 + \mathbf{n}_2\colon y \leftrightarrow z] \end{aligned}$$

and the claimed bound (4.7) follows by applying the switching lemma (2.5) with $g = z$. ∎

The following is a useful consequence, obtained by applying this lemma with $B = B_v$ for some lattice site v, $z = g$, and with the couplings given by (4.5). In this context, we distinguish between the term in (4.7) having $y = g$ and the terms for which y is a lattice site. [To simplify the notation, the superscript (N) of $Q_\alpha^{(N)}$ is omitted; the entire argument takes place at constant N.]

Corollary 4.2. If f is a decreasing function of the flux numbers, then

$$\sum_{\partial \mathbf{n}_1 = \varnothing, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2)\, I[\mathbf{n}_1 + \mathbf{n}_2 \colon B_v \leftrightarrow g]$$
$$\leqslant \sum_\alpha \beta h_\alpha \sum_{\partial \mathbf{n}_1 = \{(v,\alpha), g\}, \partial \mathbf{n}_2 = \varnothing} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2)$$
$$+ \sum_{\alpha, y \neq v, \delta} \beta J_{\{(v,\alpha),(y,\delta)\}}$$
$$\times \sum_{\partial \mathbf{n}_1 = \{(v,\alpha), g\}, \partial \mathbf{n}_2 = \{(y,\delta), g\}} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, f(\mathbf{n}_1 + \mathbf{n}_2) \qquad (4.12)$$

We are ready to proceed to the proof of the main result of this section, Theorem 2(ii), which is restated below. This extension of Theorem 2(i) is proven by suitable "block" versions of the arguments used in Section 3.

Theorem 2(ii). For GS spins with the Hamiltonian (4.3),

$$M \leqslant \beta h \chi + (\beta h M^2 + \beta\, |J|\, M^3) + (\beta h M + \beta\, |J|\, M^2) \beta\, \partial M / \partial \beta \qquad (4.13)$$

Remark. The inequality for a non-translation-invariant system is

$$M_x \leqslant \sum_v \beta h_v\, \partial M_x / \partial(\beta h_v) + \left(\beta h_x M_x^2 + \beta \sum_y J_{x,y} M_x^2 M_y \right)$$
$$+ \sum_{u,v} \left(\beta h_v M_v + \beta \sum_y J_{v,y} M_v M_y \right) \beta J_{u,v}\, \partial M_x / \partial(\beta J_{u,v}) \qquad (4.14)$$

Proof. As explained above, we begin by adding two dummy summations in (4.6):

$$M = \sum_\alpha Q_\alpha \sum_{\partial \mathbf{n}_1 = \{(0,\alpha), g\}, \partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-3}\, w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3) \qquad (4.15)$$

The current configurations in (4.15) are decomposed as in (3.5), except that the condition that a site is connected to g is replaced by a block condition. That is, for each α we identify the first block B_v visited by the backbone $\Omega_{\{(0,\alpha), g\}}(\mathbf{n}_1)$ that is also connected to g by $\mathbf{n}_2 + \mathbf{n}_3$, and, as in Section 3, we distinguish three cases.

1. There is no such block. By the same argument used for the ordinary Ising model, we obtain here a contribution

$$\begin{aligned} T &\leqslant \sum_{\alpha} Q_\alpha \sum_{x,\delta} \tanh(\beta h_\alpha) \sum_{\partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-2} w(\mathbf{n}_2)\, w(\mathbf{n}_3) \\ &\quad \times \langle \sigma_{(0,\alpha)} \sigma_{(x,\delta)} \rangle_{[C(g)]^c} I[\mathbf{n}_2 + \mathbf{n}_3 \colon (0, \alpha) \not\leftrightarrow g] \\ &\leqslant \beta h \sum_{\alpha,x,\delta} Q_\alpha Q_\delta \langle \sigma_{(0,\alpha)} ; \sigma_{(x,\delta)} \rangle \\ &= \beta h \chi \end{aligned} \tag{4.16}$$

2. The first such block contains the origin, i.e., $v = 0$. The corresponding contribution is handled by applying (4.12) to the current configurations $\mathbf{n}_2$ and $\mathbf{n}_3$ and then using (4.2) to rewrite the resulting expression in terms of the macroscopic spins:

$$\begin{aligned} R_0 &= \sum_{\alpha} Q_\alpha \sum_{\partial \mathbf{n}_1 = \{(0,\alpha), g\}, \partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-3} w(\mathbf{n}_1)\, w(\mathbf{n}_2)\, w(\mathbf{n}_3) \\ &\quad \times I[\mathbf{n}_2 + \mathbf{n}_3 \colon B_0 \leftrightarrow g] \\ &\leqslant \beta h \langle \varphi_0 \rangle^2 + \sum_{y} \beta J_y \langle \varphi_0 \rangle^2 \langle \varphi_y \rangle \\ &= \beta h M^2 + \beta\, |J|\, M^3 \end{aligned} \tag{4.17}$$

3. The first such block B_v corresponds to $v \neq 0$. The contribution of these configurations can be written as

$$\sum_{\alpha,u,\delta,v,\gamma} R_{\{\alpha,(u,\delta),(v,\gamma)\}}$$

where the label $\{\alpha, (u, \delta), (v, \gamma)\}$ indicates that $((u, \delta), (v, \gamma))$ is the first step of the backbone $\Omega_{\{(0,\alpha),g\}}(\mathbf{n}_1)$ with an endpoint in B_v. As with the Ising model, we change the parity of $n_{1\{(u,\delta),(v,\gamma)\}}$ from odd to even. This produces a factor $\tanh[\beta J_{\{(u,\delta),(v,\gamma)\}}] \leqslant \beta J_{\{(u,\delta),(v,\gamma)\}}$ and the appearance of a backbone formed by two walks $\omega_1 \colon (0, \alpha) \to (u, \delta)$ and $\omega_2 \colon (v, \gamma) \to g$ with ω_1 contained in $C_{\mathbf{n}_2 + \mathbf{n}_3}(g)^c$. Thus, (2.14) [and then (2.15), (2.17), and Griffiths II; see the treatment of (3.13)] may be used to obtain the following generalization of the inequality in (3.14):

$$\begin{aligned} R_{\{\alpha,(u,\delta),(v,\gamma)\}} &\leqslant Q_\alpha \beta J_{\{(u,\delta),(v,\gamma)\}} \langle \sigma_{(v,\gamma)} \rangle \sum_{\partial \mathbf{n}_2 = \varnothing, \partial \mathbf{n}_3 = \varnothing} Z^{-2} w(\mathbf{n}_2)\, w(\mathbf{n}_3) \\ &\quad \times \langle \sigma_{(0,\alpha)} \sigma_{(u,\delta)} \rangle_{[C(g)]^c} I[\mathbf{n}_2 + \mathbf{n}_3 \colon (0, \alpha) \not\leftrightarrow g] \\ &\quad \times I[\mathbf{n}_2 + \mathbf{n}_3 \colon B_x \leftrightarrow g] \end{aligned} \tag{4.18}$$

The remaining steps exactly parallel those taken in the corresponding part of the proof of Proposition 3.1, except that instead of the switching lemma (2.5), we now use its "block version" (4.12) in reproducing the analogue of the equality in (3.14). At this point we remark that by Griffiths II the function

$$f(\mathbf{n}_2+\mathbf{n}_3)=\langle\sigma_A\rangle_{[C(g)]^c} I[\mathbf{n}_2+\mathbf{n}_3\colon (0,\alpha)\not\leftrightarrow g]$$

[where $C(g)=C_{\mathbf{n}_2+\mathbf{n}_3}(g)$] is a *decreasing* function of the flux numbers. The result of the application of (4.12) is

$$\begin{aligned} R_{\{\alpha,(u,\delta),(v,\gamma)\}} \leqslant{} & Q_\alpha \beta J_{\{(u,\delta),(v,\gamma)\}} \langle\sigma_{(v,\gamma)}\rangle \\ & \times\left[\sum_\eta \beta h_\eta + \sum_{\eta, y\neq v,\xi} \beta J_{\{(v,\gamma),(y,\xi)\}} \langle\sigma_{(y,\xi)}\rangle\right] \\ & \times \sum_{\partial\mathbf{n}_2=\{(0,\alpha),(u,\delta)\},\partial\mathbf{n}_3=\varnothing} Z^{-2}\, w(\mathbf{n}_2)\, w(\mathbf{n}_3) \\ & \times \langle\sigma_{(v,\eta)}\rangle_{[C(0,\alpha)]^c} I[\mathbf{n}_2+\mathbf{n}_3\colon (0,\alpha)\not\leftrightarrow g] \end{aligned}$$

which implies

$$\begin{aligned} \sum_{\alpha,u,\delta,v,\gamma} R_{\{\alpha,(u,\delta),(v,\gamma)\}} \leqslant{} & \tfrac{1}{2}\beta \sum_{u,v} J_{v-u} \langle\varphi_v\rangle \\ & \times \left(\beta h + \beta \sum_y J_{v-y} \langle\varphi_y\rangle\right) \langle\varphi_0;\varphi_u\varphi_v\rangle \\ ={} & (\beta h M + \beta\,|J|\, M^2)\beta\, \partial M/\partial\beta \end{aligned} \tag{4.19}$$

The inequality (4.13) is obtained by adding (4.16), (4.17), and (4.19). ∎

5. CONSEQUENCES AND EXTENSIONS OF THE DIFFERENTIAL INEQUALITIES

In this section we discuss the implications of inequalities (1.9) and (1.10)—proven as (3.1) and (4.13)—on the phase structure and the critical behavior of the models. Some extensions of the inequalities are also mentioned—in particular, models with periodic couplings are treated.

5.1. Proof of the Main Result

The bound (1.9) [and similarly, (1.10)] is a partial differential inequality (PDI) in two variables, β and βh, which is used in different ways for the study of three different regimes. Through an appropriate elimination of either $\partial M/\partial\beta$ or $\partial M/\partial(\beta h)$ from (1.9), we obtain two ordinary differential inequalities (ODIs); one provides information along the isotherms $\{\beta = \text{const}\}$ bounding the critical exponent δ, and the other applies to the ray $\{h=0, \beta > \beta_c\}$, where it yields a bound on $\hat{\beta}$. For the sharpness of the phase transition we make full use of the two-variable structure of the PDI.

To motivate the analysis, we first define the critical points β_c and β_m (for spontaneous magnetization) to be the endpoints of the high- and low-temperature regimes:

$$\beta_c = \sup\left\{\beta: \sum_x \langle\sigma_0\sigma_x\rangle_{\beta,0} < \infty\right\} \tag{5.1}$$

and

$$\beta_m = \inf\{\beta: M(\beta, 0) > 0\} \tag{5.2}$$

Since

$$\beta_c \leqslant \beta_m \tag{5.3}$$

it suffices to show that $M(\beta, 0)$ is positive for every $\beta > \beta_c$ to prove the equality of these two critical inverse temperatures—and part (i) of Theorem 1.

The first of the two ODIs mentioned above is obtained, in the Ising spin case, by combining (1.9) and (1.12) to eliminate the β-derivative of M:

$$\begin{aligned} M &\leqslant \beta h\, \partial M/\partial(\beta h) + M^3 + \beta M^2\, \partial M/\partial\beta \\ &\leqslant \beta h\, \partial M/\partial(\beta h) + M^3 + \beta\, |J|\, M^3\, \partial M/\partial(\beta h) \end{aligned} \tag{5.4}$$

[We have also simplified by using $\tanh(\beta h) \leqslant \beta h$.] As we shall see next, counting powers in (5.4) suggests that $M(\beta_c, h) \geqslant h^{1/3}$.

To indicate the application of (5.4), let us make the assumption, which will not be required in the full analysis, that, for a given β, $M(\beta, h)$ displays a strict power law behavior in the sense that

$$M = c(\beta h)^s \tag{5.5}$$

for some constant c. Then $\beta h\, \partial/\partial(\beta h)$ acts on M simply as multiplication by s. For $\beta < \beta_c$, $M(\beta, 0) = 0$ and $\chi(\beta, 0) < \infty$; hence $s = 1$ and (5.4) is of no interest. However, for $\beta = \beta_c$, M is expected to have $s = 1/\delta < 1$. Now $\beta h\, \partial M/\partial(\beta h) = (1/\delta)M$ and the left-hand side of (5.4) is no longer completely canceled by the first term on the right-hand side. So under the strict power law assumption (5.5) at $\beta = \beta_c$, we have from (5.4) that

$$\begin{aligned}(1 - 1/\delta)M - M^3 &\leqslant \beta_c \, |J| \, M^3 \, \partial M/\partial(\beta h) \\ &\leqslant \beta_c \, |J| \, M^4/\beta h \qquad (5.6)\end{aligned}$$

where (1.13) has been used in the last step. Inequality (5.6) implies that $M^3/h \geqslant \text{const} > 0$, i.e., $\delta \geqslant 3$. As we have already mentioned, this bound was first derived by Fröhlich and Sokal.[4] Note that in the mean field approximation (and in high dimensions) $\delta = 3$ for Ising models.

By the convexity of $M(\beta, h)$ in the second variable (which follows from the Griffiths inequality)

$$[\beta h/M(\beta, h)] \, \partial M/\partial(\beta h) \leqslant [M(\beta, h) - M(\beta, 0)]/M(\beta, 0)$$

which vanishes as $h \to 0$ when $\beta > \beta_m$. The PDI (1.9) also implies therefore the following lower bound on $\partial M^2/\partial \beta$ for $\beta > \beta_m$:

$$\partial M^2/\partial \beta \,|_{\beta, h=0} \geqslant (2\beta)^{-1} [1 - M^2(\beta, 0)] \qquad (5.7\text{a})$$

and using the power law assumption at $\beta = \beta_c$,

$$\partial M^2/\partial \beta \,|_{\beta = \beta_c, h=0} \geqslant (2\beta_c)^{-1} [1 - 1/\delta - M^2(\beta_c, 0)] \qquad (5.7\text{b})$$

The ODI (5.7b) indicates that $M(\beta, 0)$ is positive for every $\beta > \beta_c$, i.e., $\beta_c = \beta_m$, while (5.7a) proves the bound $\hat{\beta} \leqslant 1/2$. Both results were previously derived only for "regular" models.[1]

It turns out that the assumption (5.5) is in fact not necessary. A similar situation was encountered in Ref. 2, where the following theorem was proven.

Theorem 5.1. Let $\{M_L(\beta, \hat{h})\}$ be a sequence of positive functions defined for $\beta, \hat{h} > 0$, increasing and differentiable in both β and $\hat{h}$, and converging as $L \to \infty$ to the function $M(\beta, \hat{h})$, which is extended to $\hat{h} = 0$ so as to be continuous in $\hat{h}$ there. Suppose that the functions M_L obey

$$M_L \leqslant \hat{h} \, \partial M_L/\partial \hat{h} + M_L f(M_L) + a_1 M_L^\theta \, \partial M_L/\partial \beta \qquad (5.8)$$

and

$$\partial M_L/\beta \leqslant a_2 M_L \, \partial M_L/\partial \hat{h} \qquad (5.9)$$

where a_1, a_2, $\theta \in (0, \infty)$ and f is a (nonnegative) continuous function satisfying

(i) $f(M) \to 0$ as $M \downarrow 0$.

(ii) $\int_{[0,1]} f(M)/M \, dM < \infty$.

If there exists a β_0 for which

$$M(\beta_0, \hat{h})/\hat{h} \to \infty \tag{5.10}$$

as $\hat{h} \downarrow 0$, then when $\hat{h}$ is small,

$$M(\beta_0, \hat{h}) \geqslant C_1 \hat{h}^{1/(1+\theta)} \tag{5.11}$$

and furthermore, for each $\beta > \beta_0$,

$$M(\beta, 0) \geqslant C_2(\beta - \beta_0)^{1/\theta} \tag{5.12}$$

with two positive constants C_1 and C_2.

Proof of Theorem 1. To see that Theorem 5.1 implies Theorem 1 for Ising models, we take $\beta_0 = \beta_c$ and observe that (5.8), with $\theta = 2$, and (5.9) are satisfied by (1.9) and (1.12). The condition (5.10), i.e., the divergence of $M(\beta_c, \beta_c h)/\beta_c h$ as $h \downarrow 0$, follows from (1.13) and the fact that $\chi(\beta_c, \hat{h})$ diverges as $\hat{h} \downarrow 0$ [if $M(\beta_c, 0) = 0$; otherwise there is nothing to prove]. As explained above, with $\theta = 2$, (5.11) provides the bound $\delta \geqslant 3$ and part (iii) of Theorem 1, and (5.12) yields both the bound $\beta \leqslant 1/2$ [part (ii) of Theorem 1] and the identity $\beta_c = \beta_m$, which is the first part of that theorem.

For the other models having spin measures in the GS class a little more work is necessary to turn (1.10) into an inequality that may be used in Theorem 5.1 with $\theta = 2$. What is needed is to replace some of the terms βh by a quantity proportional to M. Such a bound is provided by Lemma 5.2 below. Substituting the bound on $\beta h/M$ that it provides into (1.10), one gets the following inequality for general GS spins with sufficiently small βh:

$$M \leqslant \beta h\, \partial M/\partial \beta h + (\beta + 2\beta_{\mathrm{MF}})\, |J|\, M^2(\partial M/\partial \beta + M) \tag{5.13}$$

where β_{MF} is the mean field value for β_c, defined below. Inequality (5.13) shows that (5.8) is satisfied for the GS models with $\theta = 2$, and so Theorem 5.1 also implies Theorem 1 for these models. ∎

Here is the bound we used in the treatment of GS spins in the last argument.

Lemma 5.2. For a given single spin distribution ρ_0, let $S^2 = \int \varphi^2 \rho_0(d\varphi)$. Then for all h such that $\beta h S \leqslant \varepsilon$,

$$\beta h \leqslant [(\varepsilon/\tanh \varepsilon)\, \beta_{\mathrm{MF}}/|J|\,] M \tag{5.14}$$

with $\beta_{\mathrm{MF}} = (|J|\, S^2)^{-1}$.

Proof. By Griffiths II, it suffices to prove (5.14) with M replaced by the single site magnetization $\tilde{M} = \langle \varphi \rangle_{J\equiv 0}$. The integration of

$$\partial \tilde{M}/\partial(\beta h) = \langle \varphi^2 \rangle_{J\equiv 0} - \langle \varphi \rangle_{J\equiv 0}^2 \geqslant S^2 - \tilde{M}^2$$

yields the mean field bound

$$\tilde{M}/S \geqslant \tanh(\beta h S)$$

Since $(\tanh x)/x$ is a decreasing function of x, it follows that

$$\tilde{M}/S \geqslant \beta h S[(\tanh \varepsilon)/\varepsilon]$$

for every $\varepsilon \geqslant \beta h S$, which implies (5.14) for the full magnetization. ∎

5.2. Periodic and Weakly Inhomogeneous Systems

We remark that inequalities (1.9), (1.10), and (1.12) have interesting extensions to the inhomogeneous case where the coupling constants $J_{x,y}$ cease to be translation-invariant. In the case of the Ising model, inequality (3.2)—the inhomogeneous counterpart to (1.9)—implies that

$$M_x \leqslant \beta h \chi_x + M_x^3 + \beta \hat{M}^2 \, \partial M_x/\partial \beta \tag{5.15}$$

where

$$M_x = M_x(\beta, h) = \langle \sigma_x \rangle, \qquad \hat{M} = \sup_y M_y, \qquad \chi_x = \partial M_x/\partial h$$

The inhomogeneous version of (1.12) is

$$\partial M_x/\partial \beta \leqslant |J| \, \hat{M} \chi_x \tag{5.16}$$

where

$$|J| = \sup_y \left\{ \sum_z J_{y,z} \right\}$$

The GS equivalent of (5.15) can be read from (4.14).

We say a model is only *weakly inhomogeneous* if there exists a constant C ($<\infty$) so that $M_0 \leqslant C M_x$ for every x that may be connected to the origin by a path of bonds having nonzero couplings J_b. It can be easily checked that the analysis used above to prove the equality of the critical points β_c and β_m as well as the critical exponent bounds $\delta \geqslant 3$ and $\hat{\beta} \leqslant 1/2$ applies more generally to the class of weakly inhomogeneous models. This class includes the important case of *periodic* Hamiltonians.

5.3. A General Inequality

Finally, we mention that the derivation of the inequality (1.9) also leads to the following Ising model bound for general lattices (where the ghost is treated as just another site):

$$G(0, x) \leqslant G^3(0, x) + \sum_{u:u \neq x} \sum_{v:v \neq 0,u} \tanh(\beta J_{u,v})\, G^2(v, x) \times \partial G(0, x)/\partial(\beta J_{u,v}) \tag{5.17}$$

with $G(0, x) = \langle \sigma_0 \sigma_x \rangle$. The case on which we focused in this paper was that in which x is the ghost site for an otherwise translation-invariant lattice. In this situation, the sum in (5.17) should be broken into two parts: one corresponding to $v = g$, which leads to the $\partial M/\partial \beta h$ term in (1.9), and the other (where v ranges over all lattice sites) differing from the $\partial M/\partial \beta$ term in (1.9) only by an insignificant factor of 2.

Inequality (5.17), which contains a hint of the φ^4 structure present in Ising models, may prove to be of use also in the quite different context where x varies over the regular lattice sites, in which case it provides a statement about the correlation functions. However, the implications of this statement are not clear to us at this point.

To demonstrate the potential for a useful result to pass unexploited, let us mention here that (5.17) is a special case of the following bound on

$$u_4(x_1, x_2, x_3, x_4) = \langle \sigma_{x_1} \sigma_{x_2} \sigma_{x_3} \sigma_{x_4} \rangle - \langle \sigma_{x_1} \sigma_{x_2} \rangle \langle \sigma_{x_3} \sigma_{x_4} \rangle - \langle \sigma_{x_1} \sigma_{x_3} \rangle \langle \sigma_{x_2} \sigma_{x_4} \rangle - \langle \sigma_{x_1} \sigma_{x_4} \rangle \langle \sigma_{x_2} \sigma_{x_3} \rangle$$

found already in Ref. 3, Proposition 4.1:

$$\begin{aligned} |u_4(x_1, x_2, x_3, x_4)| \leqslant \sum_{u,v} \tanh(\beta J_{u,v})\, G(x_4, v)\, G(x_2, v) &\times \partial G(x_1, x_3)/\partial(\beta J_{u,v}) \\ &+ G(x_4, x_1)\, G(x_3, x_1)\, G(x_2, x_1) \\ &+ G(x_4, x_3)\, G(x_2, x_3)\, G(x_1, x_3) \end{aligned} \tag{5.18}$$

(with the restriction $u \neq x$ appearing in the proof). Inequality (5.18) was stated in Ref. 3 as a general relation (i.e., not restricted to the homogeneous setup); however, its implication for the one-point function was not appreciated there. The main concerns of that work were logarithmic bounds on the susceptibility and their implications for the renormalized coupling constant in four-dimensional models.

With $x_1 = 0$ and $x_2 = x_3 = x_4 = g$, $u_4(x_1, x_2, x_3, x_4)$ reduces to $-2G(0, x)$ and (5.18) yields (5.17). As mentioned above, (5.18) was weakened in its statement by relaxing the constraints on the sites u and v. For the main application in Ref. 3 this was an insignificant difference, but for us it is crucial to retain the fact that $u \neq x$. Otherwise the term leading to $\partial M/\partial(\beta h)$ would enter with an extra factor of 2, which for our purpose would be a nontrivial weakening of the inequality (5.17).

ACKNOWLEDGMENTS

This research was supported in part by NSF grant PHY-8605164.

REFERENCES

1. M. Aizenman, Rigorous studies of critical behavior. II, in *Statistical Physics and Dynamical Systems: Rigorous Results*, J. Fritz, A. Jaffe, and D. Szász, eds. (Birkhäuser, Boston, 1985); Absence of an intermediate phase for a general class of one-component ferromagnetic models, *Phys. Rev. Lett.* **54**:839–842 (1985).
2. M. Aizenman and D. J. Barsky, Sharpness of the phase transition in percolation models, *Commun. Math. Phys.* **108**:489–526 (1987).
3. M. Aizenman and R. Graham, On the renormalized coupling constant and the susceptibility in φ_4^4 field theory and the Ising model in four dimensions, *Nucl. Phys. B* **225**[FS9]:261–288 (1983).
4. J. Fröhlich and A. D. Sokal, The random walk representation of classical spin systems and correlation inequalities. III. Nonzero magnetic field, in preparation.
5. J. T. Chayes and L. Chayes, An inequality for the infinite cluster density in Bernoulli percolation, *Phys. Rev. Lett.* **56**:1619–1622 (1986).
6. R. B. Griffiths, Rigorous results for Ising ferromagnets of arbitrary spin, *J. Math. Phys.* **10**:1559–1565 (1969).
7. B. Simon and R. B. Griffiths, The $(\varphi^4)_2$ field theory as a classical Ising model, *Commun. Math. Phys.* **33**:145–164 (1973).
8. T. D. Lee and C. N. Yang, Statistical theory of equations of state and phase transitions, II, Lattice gas and Ising model, *Phys. Rev.* **87**:410–419 (1952).
9. R. Peierls, On Ising's model of ferromagnetism, *Proc. Camb. Phil. Soc.* **32**:477–481 (1936); see also R. B. Griffiths, Peierls proof of spontaneous magnetization in a two-dimensional Ising ferromagnet, *Phys. Rev.* **136A**:437–439 (1964); R. L. Dobrushin, The existence of a phase transition in the two- and three-dimensional Ising models, *Teorija Verojatn. Prim.* **10**:209–230 (1965).
10. P. J. Thouless, Long range order in one-dimensional Ising systems, *Phys. Rev.* **187**:732–733 (1969); see also F. Dyson, Existence of a phase transition in a one-dimensional Ising ferromagnet, *Commun. Math. Phys.* **12**:91–107 (1969); P. W. Anderson, G. Yuvall, and D. R. Hamaan, Exact results in the Kondo problem. II. Scaling theory, qualitatively correct solution and some new results on one-dimensional classical statistical models, *Phys. Rev. B* **1**:4464–4473 (1970).
11. J. Fröhlich and T. Spencer, The phase transition in the one-dimensional Ising model with the $1/r^2$ interaction energy, *Commun. Math. Phys.* **84**:87–101 (1982); see also M. Aizenman, J. T. Chayes, L. Chayes, and C. M. Newman: Discontinuity of the order parameter in one dimensional $1/|x-y|^2$ Ising and Potts models, to be submitted for publication.

12. J. M. Kosterlitz and D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C* **6**:1181–1203 (1973); see also S. Elitzur, R. Pearson, and J. Shigemitsu, Phase structure of discrete Abelian spin and gauge systems, *Phys. Rev. D* **19**:3698–3714 (1979).
13. J. Fröhlich and T. Spencer, The Kosterlitz–Thouless transition in two-dimensional Abelian spin systems and the Coulomb gas, *Commun. Math. Phys.* **81**:527–602 (1981).
14. B. Simon, Correlation inequalities and the decay of correlations in ferromagnets, *Commun. Math. Phys.* **77**:111–126 (1980).
15. J. L. Lebowitz, Coexistence of phases in Ising ferromagnets, *J. Stat. Phys.* **16**:463–476 (1977).
16. M. Aizenman and R. Fernández, On the critical behavior of the magnetization in high-dimensional Ising models, *J. Stat. Phys.* **44**:393–454 (1986).
17. K. G. Wilson and M. E. Fisher, Critical exponents in 3.99 dimensions, *Phys. Rev. Lett.* **28**:240–243 (1970).
18. A. B. Harris, T. C. Lubensky, W. K. Holcomb, and C. Dasgupta, Renormalization group approach to percolation problems, *Phys. Rev. Lett.* **35**:327–330 (1975).
19. M. Aizenman, Geometric analysis of φ_4^2 fields and Ising models. Parts I and II, *Commun. Math. Phys.* **86**:1–48 (1982).
20. M. Aizenman and C. M. Newman, Tree graph inequalities and critical behavior in percolation models, *J. Stat. Phys.* **36**:107–143 (1984).
21. R. B. Griffiths, C. A. Hurst, and S. Sherman, Concavity of the magnetization of an Ising ferromagnet in a positive external field, *J. Math. Phys.* **11**:790–795 (1970).
22. M. Aizenman, Rigorous studies of critical behavior, in *Applications of Field Theory in Statistical Mechanics*, L. Garrido, ed. (Springer-Verlag, Lecture Notes in Physics, 1985).
23. R. B. Griffiths, Correlations in Ising ferromagnets. II. External magnetic fields, *J. Math. Phys.* **8**:484–489 (1967).
24. D. C. Brydges, J. Fröhlich, and T. Spencer, The random walk representation of classical spin systems and correlation inequalities, *Commun. Math. Phys.* **83**:123–150 (1982); see also J. Fröhlich, On the triviality of $\lambda\phi_d^4$ theories and the approach to the critical point in $d \underset{(-)}{>} 4$ dimensions, *Nucl. Phys. B* **200**[FS4]:281–296 (1982); D. Brydges, J. Fröhlich, and A. D. Sokal, The random walk representation of classical spin systems and correlation inequalities. II. The skeleton inequalities, *Commun. Math. Phys.* **91**:117–139 (1983).