Journal of Statistical Physics, Vol. 33, No. 2, 1983

Finite-Size Effects at First-Order Transitions

Vladimir Privman[1] **and Michael E. Fisher**[1]

Received May 24, 1983

Finite-size rounding of a first-order phase transition is studied in "block"- and "cylinder"-shaped ferromagnetic scalar spin systems. Crossover in shape is investigated and the universal form of the rounded susceptibility peak is obtained. Scaling forms on the low-temperature side of the critical point are considered both above and below the borderline dimensionality, $d_{>} = 4$. A method of phenomenological renormalization, applicable to both odd and even field derivatives, is suggested and used to estimate universal amplitudes for two-dimensional Ising models at $T = T_c$.

KEY WORDS: First-order transitions; finite-size effects; scaling theory; Ising models; phenomenological renormalization; borderline dimensionality.

1. INTRODUCTION

In this paper we study various finite-size effects arising at the zero-field phase boundary of a ferromagnetic scalar (or Ising-like) spin system, with an emphasis on the rounding of the first order transition for $T < T_c$. Finite-size scaling theory[1,2] for behavior close to a critical point has attracted appreciable theoretical and experimental effort (see Ref. 3 for a recent review). Transfer matrix calculations utilizing finite-size scaling ideas have led to some remarkably precise and accurate estimates for the critical exponents of several ($d = 2$)-dimensional models (see the overview presented by Nightingale[4]). More recently, a study of finite size effects at *first-order* phase transitions has been initiated. A scaling theory was developed,[5] and the properties of renormalization groups close to the associated "discontinuity" fixed points[6] were invoked. It was found[5,7–9] that

[1] Baker Laboratory, Cornell University, Ithaca, New York 14853.

0022-4715/83/1100-0385$03.00/0

the rounding of a first-order transition as a function of the ordering field, say, H, takes place on a scale proportional to the inverse of the total volume, V, of the system. However, if the system is already infinite in one direction, so forming a cylinder of cross-sectional area A, the rounding in field is known to be *exponentially* narrow in A; this property has been derived analytically[9-11] for the field-driven transitions in Ising model cylinders and numerically[9,12] for temperature-driven first-order transitions in several ($d = 2$)-dimensional Potts models with $q > 4$ states per site. One of the motivations for our present work is to understand the crossover in the sharpness of a first-order transition that must evidently take place when the shape of a system goes over from a totally finite "block" geometry to an elongated, cylindrical geometry with, ultimately, one infinite dimension. In addition we hoped to obtain more explicit forms for the scaling functions describing the characteristic rounding behavior.

In Section 2 we combine an introductory discussion with a derivation of the rounding form for the "block" situation and we present a scaling formulation which allows a crossover to cylindrical geometry. In Section 3 we call on the transfer matrix approach to obtain an explicit form for the rounding which interpolates between the block and cylinder geometries. We compare our results with an approximate scaling analysis for low temperatures presented recently by Cardy and Nightingale[13] (see also Appendix B).

Section 4 addresses various further questions including the nature of the "corrections-to-scaling" for different types of boundary condition. We also exhibit the rounding behavior in the solvable infinite-range model which displays mean field critical behavior. Cylindrical geometry is taken up again and two regimes of the exponential rounding are identified: one close to the $d = 1$, $T \equiv 0$, limiting phase transition, and the other arising when the d-dimensional phase transition is approached at fixed temperature in the range $0 < T < T_c$. The interplay between the two forms of behavior is related to the detailed nature of the asymptotic degeneracy of the two largest eigenvalues of the transfer matrix. For the standard square lattice Ising model, we report (in Appendix A) some exact results, based on Onsager's solution.[14,15]

Recently Brézin[16] has argued on the basis of exact calculations in the infinite-component, $n \to \infty$ limit, that the usual finite-size scaling theory for the *critical* region[1,2] is valid only for d less than the borderline dimensionality $d_> = 4$. In Section 5 we consider the approach to criticality and derive the large argument asymptotics of the scaling functions for both block and cylinder geometries. We present a generalization of the finite-size scaling hypothesis which is applicable for $d > d_>$ and which agrees with Brézin's explicit results: a "dangerous irrelevant variable" plays a central role.

Finally, in Section 6, we argue that the so-called "single-phase functions" introduced by Schulman and Privman[10,11] satisfy a finite-size scaling postulate near criticality. This property allows phenomenological renormalization calculations[4] to be performed for both even and *odd* field derivatives of the free energy below and at T_c. We illustrate the application of the method and report numerical estimates for the first few critical exponents and universal critical amplitudes for the square and the triangular lattice Ising models. Some open problems are mentioned briefly in Section 7.

2. SCALING FORMS FOR FIRST-ORDER TRANSITIONS

For definiteness we consider d-dimensional hypercubic lattices with lattice spacing a and cell volume a^d. A rectangular or *block* lattice geometry will be specified by sides of lengths $L_1, L_2, \ldots L_d$ parallel to the lattice axes. Frequently, we will consider the simple, finite geometry

$$L_1 = L_{\|} \equiv L, \qquad L_2 = L_3 = \cdots L_d = L_{\perp} \tag{2.1}$$

which represents a *rod* if $L_{\|} > L_{\perp}$, as we will usually consider, or a *slab* if $L_{\|} < L_{\perp}$. The total volume is

$$V = \prod_{i=1}^{d} L_i = L_{\perp}^{d-1} L_{\|} \tag{2.2}$$

and the cross-sectional areas are

$$A_j = V/L_j = \prod_{i \neq j}^{d} L_i \qquad \text{with} \quad A \equiv A_1 = L_{\perp}^{d-1} \tag{2.3}$$

Unless explicitly stated otherwise, *periodic boundary conditions* will be assumed in each direction. An infinite *cylinder* geometry is specified by $L_1 = L_{\|} \equiv L \to \infty$.

At each site i, with position vector $\mathbf{r}_i$, we suppose a *scalar* spin variable, s_i, is located which interacts with an external magnetic field

$$H = h k_B T \tag{2.4}$$

and, via couplings of finite range, with spins on other sites, j. Various explicit calculations will be discussed and presented for the nearest neighbor spin-1/2 ferromagnetic Ising model with Hamiltonian

$$\mathscr{H} = -J \sum_{\langle ij \rangle} s_i s_j - H a^d \sum_i s_i \qquad (s_i = \pm 1) \tag{2.5}$$

the unorthodox factor a^d is introduced here for dimensional reasons that will become more evident below. For simplicity we will use the language of this model even when our considerations can be readily generalized.

The scaling ansatz for first-order transitions advanced by Fisher and Berker (FB)[5] can be written for the singular part of the reduced free energy density of a simple ferromagnet below criticality, $T < T_c$, in the form

$$f_s(H, T; L_j) \equiv F_s/k_B TV \approx A_0(T) L_0^{-d} W\left[B(T) H L_0^d; l_j\right] \tag{2.6}$$

where the shape ratios $l_j = L_j/L_0$ are assumed fixed and of order unity: a convenient normalization is provided by

$$\prod_{j=1}^{d} l_j = 1 \quad \text{so that} \quad L_0^d = V \tag{2.7}$$

Technically the temperature is an irrelevant variable when $T < T_c$ and the coefficients $A_0(T)$ and $B(T)$ then represent nonuniversal amplitudes. The scaling function, $W(y; l_j)$, must, when conveniently normalized, satisfy

$$W(y; l_j) \approx -|y|, \qquad \text{as} \quad y \to \pm\infty \tag{2.8}$$

in order to reproduce the first-order transition in the infinite-size limit, $L_0 \to \infty$. Indeed, the magnetization density in that limit becomes

$$m \equiv -\frac{1}{V}\left(\frac{\partial F_s}{\partial H}\right)_T \approx \pm k_B T A_0(T) B(T) = \pm m_0(T) \tag{2.9}$$

for small $H \gtrless 0$, where $m_0(T)$ is the bulk spontaneous magnetization which vanishes near criticality as

$$m_0(T) \sim |t|^\beta \qquad \text{with} \quad t = (T - T_c)/T_c \tag{2.10}$$

The zero-field susceptibility in the finite system is then given by

$$\chi_0(T, L_0) = \left(\frac{\partial m}{\partial H}\right)_{H=0} \approx -k_B T A_0 B^2 W_0''(l_j) L_0^d \equiv C(T; l_j) V \tag{2.11}$$

where $W_0''(l_j) = (d^2W/dy^2)_{y=0}$. Now, as discussed in FB, the zero-field susceptibility for large L_0 can be expressed in terms of the correlation functions, $\langle s_i s_j \rangle$, for large spin separations. Specifically, the fluctuation relation yields

$$\chi_0 = (a^{2d}/Vk_B T) \sum_i \sum_j \langle s_i s_j \rangle \tag{2.12}$$

The zero-field correlation functions, $\langle s_i s_j \rangle$, depend here, of course, on the size and shape of the system but for $T < T_c$ we may argue that the predominant configurations of the system correspond to the spins being spontaneously magnetized "up" or spontaneously magnetized "down." This encapsulates the usual argument[5,17–19] leading to the identification of the (*short*) long-range order as

$$\lim_{|\mathbf{r}_i - \mathbf{r}_j| \to \infty} \lim_{L_0 \to \infty} \langle s_i s_j \rangle = m_0^2(T) \tag{2.13}$$

and to the conclusion that when $L_0 \to \infty$ the sum in (2.12) may be correctly estimated by the corresponding replacement $\langle s_i s_j \rangle \Rightarrow m_0^2$.[17-19] If, at least provisionally, we accept this heuristic argument we find, for $L_0 \to \infty$,

$$\chi_0 \approx (m_0 V)^2 / k_B T V \quad \text{or} \quad C(T; l_j) = m_0^2 / k_B T \tag{2.14}$$

which, although nonrigorous, would seem to be rather generally valid.

Combining (2.9), (2.11), and (2.14) yields the identifications

$$A_0 = -W_0''(l_j), \qquad B(T) = m_0(T)/k_B T A_0 \tag{2.15}$$

from which we conclude that A_0 is independent of T and that $W_0''(l_j)$ is actually independent of the shape ratios, l_j (see also below). More importantly, we see from (2.6) that the finite size of the system at the first-order transition enters the FB scaling ansatz *only* through the natural combination

$$y_V = E_V(T)/k_B T = m_0(T) H V / k_B T \tag{2.16}$$

which represents the dimensionless ratio of the total bulk ordering energy of the transition to the thermal energy! This constitutes a most appealing and, as we will see, rather general way of restating the main scaling conclusion of FB.

If one takes to heart, more seriously still, the predominance of the two spontaneously magnetized configurations, the partition function for the system near the transition should be approximated well by

$$Z(H; T, L_j) \sim e^{+m_0 h V} + e^{-m_0 h V} \tag{2.17}$$

This leads immediately to the explicit results

$$f_s(H, T; L_j) \approx -V^{-1} \ln\{2 \cosh[m_0(T) H V / k_B T]\} \tag{2.18}$$

$$m_s(H, T; L_j) \approx m_0(T) \tanh[m_0(T) H V / k_B T] \tag{2.19}$$

which, clearly, are in the expected scaling forms, the scaling function in (2.6) being $W(y_V) = -\ln(2 \cosh y_V)$ while $A_0 = 1$.

The simple scaling picture thus obtained is intuitive and, seemingly, quite satisfactory. It cannot, however, be regarded as complete: in particular, it indicates that the rounding of the transition is always on the scale

$$H_V \simeq k_B T / m_0(T) V \tag{2.20}$$

which is *shape-independent*, whereas we know[9-11] that in the case of a long block or rod with $L_\parallel \gg L_\perp$, which approaches a cylinder if $L_\parallel \to \infty$, the rounding should become *exponentially* small in $L_\perp$. Likewise, χ_0 always diverges linearly with $V = \prod_j L_j$ according to (2.11) [or (2.19)], whereas for a cylinder χ_0 diverges exponentially rapidly with $L_\perp$. We aim to understand this deficiency of the simple scaling analysis and to repair it so that the crossover between block and cylinder geometry can be described more satisfactorily.

One unstated assumption leading to (2.16)–(2.20) is immediately obvious, namely, here (as in FB) it has been implicitly assumed that the dimensions satisfy

$$L_j \gg \xi_\infty(T) \qquad (\text{all } j) \tag{2.21}$$

where ξ_∞ represents the *bulk* ($L_j \to \infty$) *single-phase correlation length* measured by the decay, or by the second spatial moment of the *net* correlation function

$$G(\mathbf{r}_i - \mathbf{r}_j) = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$$

This condition is needed in order that the long-distance behavior of the correlation functions, $\langle s_i s_j \rangle$, dominates in (2.12). However, it should cause no problems except near the critical point where ξ_∞ diverges as $|t|^{-\nu}$.

More serious is the fact that in asserting the predominance of the states of "up" or "down" with total magnetization $\pm V m_0(T)$, we have overlooked all configurations in which some regions of the system are magnetized "up" while others are magnetized "down," as illustrated in Fig. 1. Such configurations are, of course, suppressed by a Boltzmann factor representing the excess free energy associated with the interface (or domain wall) between the oppositely magnetized regions: Including them does, however, increase the entropy. For *block* geometry, configurations of this sort may thus yield corrections which, relative to the bulk, will be of order $A/V \sim 1/L_0$. This already suggests corrections to (2.18) [or (2.6)] of order $1/V^{1/d}$ which remain undisplayed. (See, however, Sections 3 and 4 below.)

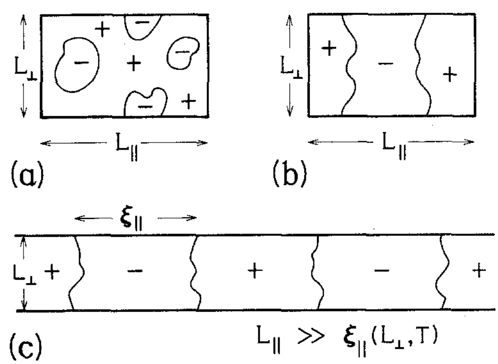

Fig. 1. Some typical configurations of a system of finite size exhibiting nonuniform ordering (or magnetization) forming "up" (or +) and "down" (or −) domains in a block geometry, (a) and (b), and in a cylinder geometry, (c), where $L_\parallel \to \infty$.

In the case of a *cylinder* geometry the effects are more serious. As illustrated in Fig. 1c, whenever $L_{\parallel} \gg L_{\perp}$ we expect the dominant configurations to involve domain walls of area $A = L_{\perp}^{d-1}$ which reach *across* the system so resulting in fluctuations which break the system into successive regions of "up" and "down" magnetization of a characteristic length which we will call $\xi_{\parallel}(T; L_{\perp})$. This length may be expressed in terms of the ratio of the two largest eigenvalues of the transfer matrix below T_c (see Section 3); but, more generally, it is related to the interfacial tension (or domain wall free energy) $\Sigma(T)$ via[15]

$$\xi_{\parallel}(T; L_{\perp}) \sim \exp[A\Sigma(T)/k_B T] \tag{2.22}$$

so diverging exponentially fast as $L_{\perp} \to \infty$. (There is a more slowly varying prefactor present here which will be discussed in Sections 4 and 5.)

These observations suggest that for a cylinder (with $L_{\parallel} = \infty$) scaling should involve the ratio

$$y_A = E_A(T)/k_B T = m_0(T) H A \xi_{\parallel}(T; A)/k_B T \tag{2.23}$$

which would replace (2.16). Thus we might expect the scaling relation

$$m_s(H, T; L_{\parallel} = \infty, L_{\perp}) \approx m_0(T) Y_{\infty}[m_0 h L_{\perp}^{d-1} \xi_{\parallel}(L_{\perp})] \tag{2.24}$$

to hold for a cylinder. This then predicts rounding on the scale

$$H_A = k_B T / m_0(T) \xi_{\parallel}(T; L_{\perp}) A \sim \exp(-\sigma L_{\perp}^{d-1}) \tag{2.25}$$

where the reduced interfacial tension is

$$\sigma(T) = \Sigma(T)/k_B T \tag{2.26}$$

Thus we obtain the anticipated exponentially small rounding and an exponential divergence of $\chi_0(L_{\perp})$. The surmise (2.24) will be justified analytically in Section 3 and an explicit form for the scaling function $Y_{\infty}(y)$ will be obtained.

The presence of a diverging correlation length, $\xi_{\parallel}$, suggests that when $L_{\parallel}$ is finite it should enter specifically through the combination $x = L_{\parallel}/\xi_{\parallel}$ which is actually the same as the scaling ratio y_V/y_A. Thus we may extend (2.6) [or (2.19) and (2.24)] to obtain the combined scaling forms for the magnetization

$$\begin{aligned} m(H, T; L_j) &\approx m_0(T) Y(y_V, y_A) \\ &\approx m_0(T) Y_0[m_0 h V; L_{\parallel}/\xi_{\parallel}(L_{\perp})] \end{aligned} \tag{2.27}$$

which should describe *both block* and *cylinder* geometries and the crossover between them. Thus when $L_{\parallel}/\xi_{\parallel}$ becomes large we should have $Y(y_V \to \infty, y_A) \to Y_{\infty}(y_A)$ so that (2.24) applies and the rounding is exponentially small. Conversely, if the shape ratios l_j remain fixed as $L_{\parallel}, L_{\perp}$

$\to \infty$ we have $L_{\|}/\xi_{\|} \to 0$ and expect

$$Y_0(y_V; x \to 0) \to \tanh y_V \tag{2.28}$$

so reproducing (2.19) and yielding rounding of order $1/V$.

Some caution is necessary, however, since if $L_{\|}$ becomes significantly smaller than $L_{\perp}$ one goes over to a slab geometry (or, for $d = 2$, simply to a rod oriented along the second axis). Since a slab of finite thickness $L_{\|}$ may exhibit a transition in the limit $L_{\perp} \to \infty$ a different crossover must arise: even in the absence of a transition, however, we should expect a different scaling combination, say, $L_{\perp}/\tilde{\xi}(T; L_{\|}, L_{\perp})$, to enter in a nontrivial way. In the limit $L_{\perp} \to \infty$ there will be finite-size corrections of order $\exp(-L_{\|}/\xi_{\infty})$ (or some power thereof) and, in the case of a first-order transition, interface corrections of order $\exp[-L_{\|} S \Sigma(T)/k_B T]$ where S is some characteristic $(d-2)$-dimensional perimeter of a heterophase fluctuation. For $d = 2$ we clearly have, by symmetry, corrections of the form $\exp(-L_{\|}\Sigma/k_B T)$ [which, via the hyperscaling relation $\Sigma \sim 1/\xi_{\infty}$, are also of the form $\exp(-L_{\|}/\xi_{\infty})$, at least, in the critical region!] and evidently $\tilde{\xi} \sim \exp(L_{\|}\Sigma/k_B T)$. Thus it is plausible generally that the new length $\tilde{\xi}$ diverges exponentially fast with $L_{\|}$. If so it would suggest that the crossover to slab behavior occurs when $L_{\|}$ is of order $\ln L_{\perp}$; we will find that the same criterion arises in the transfer matrix analysis presented in the next section.

In summary, scaling at a first-order transition in a block geometry should be controlled by the "bulk" combination, $y_V = m_0 HV/k_B T$; but if $L_{\|}$ becomes much larger than $L_{\perp}$, the further, "cross-sectional area" combination, $y_A = m_0 HA\xi_{\|}/k_B T$ enters, with $\xi_{\|}(L_{\perp})$ diverging in accord with (2.22); in the cylinder limit, $L_{\|} = \infty$, only y_A determines the rounding of the first-order transition. We now test these conclusions in various ways: we will find that they even extend *into* the critical region.

3. TRANSFER MATRIX ANALYSIS

The transfer matrix method for Ising spin systems is well known[20] but it is not always recalled that it is of considerable generality, applying to continuous variables and extending to continuum systems.[21,22] If $\Lambda_r = \Lambda_r(H, T; L_{\perp})$ with $r = 0, 1, 2, \ldots, R(L_{\perp})$ are the eigenvalues of the transfer matrix for a system of cross-sectional dimensions $L_{\perp}$, arranged in order of decreasing magnitude, the reduced free energy density is given by

$$f(H, T; L_{\|}, L_{\perp}) = -V^{-1} \ln\left(\sum_{r=0}^{R} \Lambda_r^{L/a} \right) \tag{3.1}$$

with $L_{\|} \equiv L$. The upper limit $R(L_{\perp})$ in general diverges as $\exp(cA/a^{d-1})$, where c is a constant and $A = L_{\perp}^{d-1}$. In the cylinder limit, $L_{\perp} \to \infty$, the first

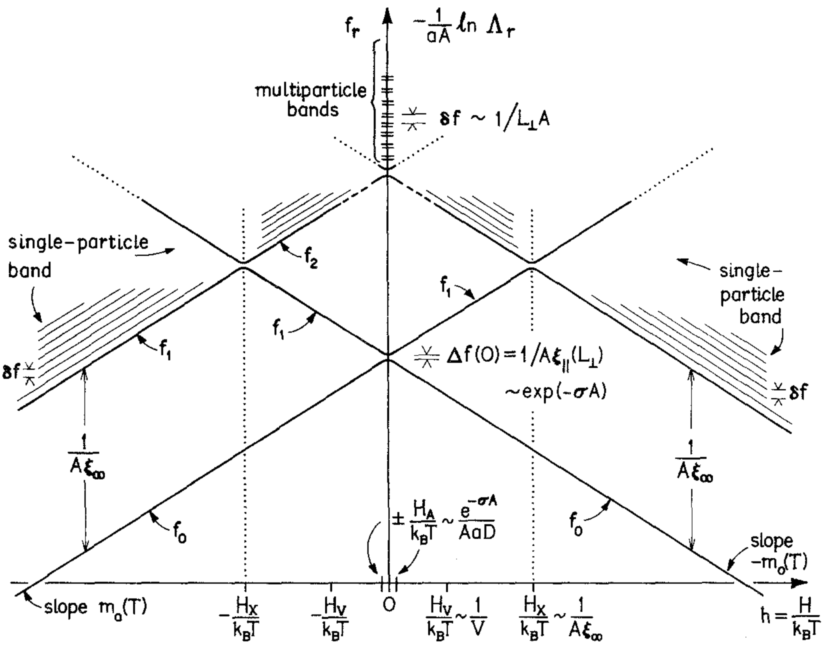

Fig. 2. Depiction of the transfer matrix spectrum for a system of cross-sectional area $A = L_\perp^{d-1}$ as a function of field, H, through a first-order transition in terms of the "free-energy levels" $f_r = -(\ln \Lambda_r)/Aa$. Details of the spectrum above f_2 are incomplete.

term in the sum gives the complete result. For $T < T_c$ and H not too large we will see, further, that the *first two terms* give a satisfactory representation[2] for a rather wide range of $L_\parallel$. To show this consider the behavior of the "free energy levels"

$$f_r(H, T; L_\perp) = -(Aa)^{-1} \ln \Lambda_r(H, T; L_\perp) \tag{3.2}$$

Our knowledge of the structure of the low-lying levels as a function of $h = H/k_B T$ is summarized schematically in Fig. 2. This figure synthesizes what is known rigorously in zero field for the two-dimensional Ising model,[(14)] what can be concluded by general analyses,[(21,22,26)] what can be surmised by studying behavior at large fields, H, and in general field for

[2] A similar conclusion regarding the adequacy of only the first two eigenvalues has been reached independently in a somewhat different context by Kleban and Akinci,[(23)] who have studied the shape dependence of the specific heat of a two-dimensional Ising model on an $m \times n$ torus in the critical region in the scaling limit, $m, n \to \infty$ with m/n fixed as analyzed originally by Ferdinand and Fisher.[(24)] Two-eigenvalue dominance in yet another context has been found by Bruce.[(25)]

$T = 0$, what can be checked in certain simple models,[27] and what can reasonably be conjectured about the nature of the asymptotic degeneracy of Λ_0 and Λ_1 and tested numerically.[10,11] The first point is that the lowest level, f_0, determines the bulk free energy and hence, for h of order unity but small, has a slope $\mp m_0(T)$. The next crucial feature concerns the asymptotic degeneracy in zero field which fixes the gap

$$\Delta f(H=0, T; L_\perp) \equiv f_1 - f_0 = 1/A\xi_\parallel(T; L_\perp) \tag{3.3}$$

in terms of the longitudinal (zero-field) correlation length,

$$\xi_\parallel(T; L_\perp)/a \approx D(T; L_\perp)\exp[A\Sigma(T)/k_B T], \qquad \text{as} \quad L_\perp \to \infty \tag{3.4}$$

and of the interfacial tension $\Sigma(T) \equiv k_B T\sigma(T)$.[15] Note we include the amplitude factor which is slowly varying in the sense that

$$A^{-1}\ln D(T; L_\perp) \to 0, \qquad \text{as} \quad L_\perp \to \infty \tag{3.5}$$

In Appendix A we outline an explicit evaluation of $D(T; L_\perp)$ for the two-dimensional Ising model and find, as might have been guessed, that it varies as a power of $L_\perp$ (see further below).

The third point is that the gap, $\Delta f = f_1 - f_0$, for $h = O(1)$ small, is equal to $1/A\xi_\infty(T)$ where, as in Section 2, ξ_∞ denotes the bulk single-phase correlation length. Above the gap lies a quasicontinuum of "single-particle" levels[21,22,26] with a spacing δf of magnitude $1/AL_\perp$ forming a band containing of order A/a^{d-1} levels. In a zeroth-order approximation (which is exact at $T = 0$) the lowest levels cross linearly in zero field and again at a field H_X which, it now follows, must be given by

$$2m_0 H_X/k_B T = 1/A\xi_\infty \tag{3.6}$$

Furthermore, it appears[10,11,27] that the "avoided crossings" that appear for $T > 0$ are, to leading order in $L_\perp$, still located at H_X. From this we can conclude that under the condition

$$T < T_c \quad \text{and} \quad |H| \ll H_X \propto k_B T/m_0\xi_\infty L_\perp^{d-1} \tag{3.7}$$

the two lowest free energy levels are separated from the higher levels by a gap of order $1/A\xi_\infty$. This in turn means that under the same conditions we have

$$\Lambda_2/\Lambda_1 < \exp(-a/\xi_\infty) \tag{3.8}$$

and hence can write the total free energy density as

$$f = -V^{-1}\ln[\Lambda_0^{L/a} + \Lambda_1^{L/a}] + \mathscr{E}(L_\parallel) \tag{3.9}$$

where $\mathscr{E}(L_\parallel)$ denotes terms exponentially small as $L_\parallel \to \infty$. To be more explicit, note that the sum $\sum_{r=2}^{R}(\Lambda_r/\Lambda_0)^{L/a}$ is, by (3.8), certainly bounded by $R \exp(-L/\xi_\infty) \approx \exp(cA/a^{d-1} - L/\xi_\infty)$ and so is exponentially small when $L \equiv L_\| \gg L_\perp^{d-1}\xi_\infty(T)/a^{d-1}$. More realistically, however, only the first, single-particle band of order $(L_\perp/a)^{d-1}$ levels need be counted to estimate the leading correction so that the error term in (3.9) is negligible when

$$L_\| \gg (d-1)\xi_\infty(T)\ln(L_\perp/a) \tag{3.10}$$

which is a much milder condition! (Compare with the penultimate paragraph in Section 2.)

Now one can argue, and substantiate by various detailed checks,[10,11,27] that the "avoided crossing" of the two largest eigenvalues as a function of the field, h, or, better, of the two lowest free energy levels, $f_0(H)$ and $f_1(H)$, may, for large $L_\perp/a$, be described correctly by the roots of a quadratic equation that represents the characteristic determinant of a real symmetric 2×2 matrix whose diagonal matrix elements are smoothly crossing, symmetrically related functions, $f_+(H)$ and $f_-(H) = f_+(-H)$, while the off-diagonal elements serve to produce the splitting $\Delta f(0)$ given by (3.3). The "single-phase free energies" thus introduced are then given by

$$f_\pm(H,T;L_\perp) = \tfrac{1}{2}(f_0+f_1) \mp \tfrac{1}{2}h\left\{(f_1-f_0)^2/h^2 - \left[\Delta f(0)/h\right]^2\right\}^{1/2} \tag{3.11}$$

Further, one can show,[10,11] as is to be anticipated, that as $L_\perp/a \to \infty$ one has

$$f_\pm \approx f_\infty(0,T) \mp m_0(T)h - \tfrac{1}{2}k_B T\chi_\infty(T)h^2 \mp O(h^3) \tag{3.12}$$

where $f_\infty(0,T) = f(H=0,T;L_j \to \infty)$ is the bulk zero-field free energy, $m_0(T)$ is, as before, the bulk spontaneous magnetization and, similarly, $\chi_\infty(T)$ is the initial susceptibility of the infinite system. The residual errors in (3.12) arise from the finite transverse dimensions of the layers: for periodic boundary conditions, we thus anticipate[1] that they are of order $\exp(-L_\perp/\xi_\infty)$, provided only that the avoided crossings between $f_1(H)$ and $f_2(H)$ at $H \simeq H_X$ are not reached: this is ensured if (3.7) is respected. Conversely, then, for the two lowest free energy levels we expect the representation

$$f_0, f_1 \approx f_\infty(T) - \tfrac{1}{2}k_B T\chi_\infty(T)h^2 \mp \left[m_0^2h^2 + 1/4A^2\xi_\|^2(L_\perp)\right]^{1/2} \tag{3.13}$$

to be accurate to order h^3 and $\exp(-L_\perp/\xi_\infty)$ as regards $f_0(H,T;L_\perp)$ although similar accuracy for $f_1(H,T;L_\perp)$ should apply only up to $|H| \lesssim H_X$ [given by (3.6)]. This expression is clearly useful in that all the rapid dependence on H has been isolated explicitly in terms of $\xi_\|(T;L_\perp)$.

If we accept (3.13) and, in order to obtain the leading crossover behavior in h, neglect the $\chi_\infty h^2$ term, we obtain from (3.2) and (3.9) the final result

$$f_s \equiv f(H, T; L_\parallel, L_\perp) - f_\infty(T)$$
$$\approx -V^{-1}\ln 2\cosh\left\{ V\left[m_0^2(T)h^2 + 1/4A^2\xi_\parallel^2(T, L_\perp)\right]^{1/2}\right\} \quad (3.14)$$

this should be valid for $T < T_c$ up to corrections in $L_\parallel$ and $L_\perp$ which are exponentially small provided that (3.10) is met and that the further condition

$$L_\perp/\xi_\infty(T) \gg 1 \quad (3.15)$$

is satisfied. Note that on combining this with (3.10) the shape ratio $L_\parallel/L_\perp$ is required to be bounded below only by $\xi_\infty \ln(L_\perp/a)/L_\perp$ which becomes arbitrarily small for large $L_\perp$; i.e., flat slabs are allowed! For finite $L_\perp$, accuracy can be improved, if desired, by replacing f_∞ in (3.14) by $\frac{1}{2}[f_+(L_\perp) + f_-(L_\perp)]$ and $m_0(T)$ by $[f_-(L_\perp) - f_+(L_\perp)]/2h$.[10,11] For systems with anisotropic interactions (3.15) must, naturally, be replaced by $L_j/\xi_\infty^{(j)} \gg 1$ where $\xi_\infty^{(j)}$ is the correlation length in the direction j.

To bring out the main features of the conclusion (3.14) consider, first, the limit of large $H/k_B T$: the result then reduces simply to

$$f_s(H, T; L_j) \approx -V^{-1}\ln 2 - m_0|h| \quad (3.16)$$

which reproduces correctly the appearance of the bulk first-order transition. Next consider the *block limit*

$$x = L_\parallel/\xi_\parallel(L_\perp) \approx L_\parallel \exp(-\sigma L_\perp^{d-1})/D(L_\perp) \to 0 \quad (3.17)$$

where we have invoked (3.4) and (3.5). Evidently one now recaptures *precisely* the naive "two-peak" result (2.18), which scales solely in terms of the bulk ratio $y_V = m_0 HV/k_B T$ with, therefore, rounding on the scale H_V [see (2.20)]. Note that the ratio of H_V to H_X is given by $2\xi_\infty/L_\parallel \ll 1/(d-1)\ln(L_\perp/a)$, where (3.10) has been recalled: thus the condition (3.7) is amply met in the region of interest.

On departure from the block limit a cursory inspection of (3.14) suggests that sharp structure in the magnetization as a function of H might set in on the scale $H_A = k_B T/m_0 A\xi_\parallel \ll H_V$ [see (2.25)]. In fact, however, this is *not* the case since when $x = L_\parallel/\xi_\parallel$ is small *and* $|H| \lesssim H_A$ the whole argument of the cosh in (3.14) is then likewise small and one may expand to obtain

$$Vf_s \approx -\ln 2 - \tfrac{1}{8}x^2 - \tfrac{1}{2}y_V^2\left(1 - \tfrac{1}{12}x^2 + \cdots\right)$$
$$+ \tfrac{1}{12}y_V^4\left(1 - \tfrac{1}{5}x^2 + \cdots\right) + O(x^4, y_V^6) \quad (3.18)$$

Consequently the deviations in magnetization, susceptibility, etc. enter merely as multiplicative factors which depart from unity as x^2.

More generally, we see that (3.14) has exactly the two-variable crossover scaling form anticipated in (2.27): explicitly we can rewrite (3.14) as

$$f_s(H, T; L_j) \approx -V^{-1}\ln\left[2\cosh\left(y_V^2 + \tfrac{1}{4}x^2\right)^{1/2}\right] \tag{3.19}$$

so that the crossover scaling function for the magnetization in (2.27) is

$$Y_0(y; x) = \frac{y}{\left(y^2 + \frac{1}{4}x^2\right)^{1/2}} \tanh\left[\left(y^2 + \tfrac{1}{4}x^2\right)^{1/2}\right] \tag{3.20}$$

which satisfies the limit relation (2.28). A similar result, with $m_0(T)$ replaced by 1 in the limit $T \to 0$ can be derived using the low-temperature approximate renormalization-cum-rescaling approach of Blöte, Nightingale and Cardy[9,13] but the form of the slowly varying prefactor, $D(T, L_\perp)$, in the expression (3.4) for $\xi_\parallel$ is not reproduced correctly: see the next section and Appendix B for details.

Lastly, notice that in the *cylinder limit* $x = L_\parallel/\xi_\parallel \to \infty$ one obtains simply

$$f_s(H, T; L_j) \approx -\left[m_0^2(T)h^2 + 1/4A^2\xi_\parallel^2(T; L_\perp)\right]^{1/2} \tag{3.21}$$

so that the scaling function for the magnetization in (2.24) becomes

$$Y_\infty(y_A) = 2y_A/\left(1 + 4y_A^2\right)^{1/2} \tag{3.22}$$

and rounding is now only on the scale $H_A \ll H_V \ll H_X$. The zero-field susceptibility hence diverges as

$$\chi_0(T; L_\perp) \approx \frac{2m_0^2(T)}{k_B T} L_\perp^{d-1}\xi_\parallel(T, L_\perp) \approx \frac{2m_0^2}{k_B T} L_\perp^{d-1} a D(L_\perp) e^{A\Sigma/k_B T} \tag{3.23}$$

where the full coefficient has now been identified.

In summary, the transfer matrix calculations bear out all the general scaling features anticipated in Section 2 and, furthermore, provide the explicit scaling functions and amplitudes describing the cylinder limit *and* the crossover from the block limit.

4. SOME FURTHER ASPECTS

The considerations of Section 2 leading to the scaling form (2.27) are, of course, not much more than heuristic; similarly, although we believe the transfer matrix analysis of Section 3 leading to the explicit result (3.14) is rather convincing, it is certainly not rigorous. The main points open to question will, we trust, have been evident to the reader. In this section we comment further on a few more detailed issues.

First, recall the restriction to periodic boundary conditions in all finite directions. Under these conditions no matrix elements are needed in (3.1) and one may assert[1-3] that corrections to the leading scaling results are of order $\exp(-L_j/\xi_\infty^{(j)})$. This conclusion should hold equally for *anti*periodic or helical boundary conditions in one or more of the transverse directions since these, also, respect the symmetry $H \Leftrightarrow -H$ and so do not shift the location of the susceptibility peak. For free or open boundary conditions, however, changes in the free energy of relative magnitude $a/L_\perp$ and $a/L_\parallel$ occur and, if there are also surface magnetic fields the position of the susceptibility peak may shift[5] to a field $H_\sigma(L_j)$ of order $(aA_i/V) \sim (a/L_i)$ which, asymptotically, is much larger than the rounding field $H_V \sim a^d/V$. One may reasonably conjecture that the *same* forms of rounding and scaling will be valid with H simply replaced by the shifted field[1] $\dot{H} = H - H_\sigma(L_j)$ but we have not investigated the more detailed arguments required to substantiate this. (Such shifts have been studied in the critical region by scaling hypotheses[28] and local mean-field theory.[29])

When the finite-width corrections are exponentially small one may, as remarked, improve the accuracy of (3.14) by using the terms in (3.12) of higher order in the field. Thus, for the susceptibility the leading correction to the scaling peak in the block limit is given, for $|H| \lesssim H_X$, by

$$\chi(H,T;L_j) \approx \frac{Vm_0^2(T)}{k_B T \cosh^2(Vm_0 h)} + \chi_\infty(T) + O(H) \tag{4.1}$$

This formula might be useful in analyzing well-equilibrated Monte Carlo data on a finite system in small fields below T_c.

It is instructive, in passing, to examine the rounding of the first-order transition in the *infinite range* or Husimi–Temperley model which yields mean-field theory in the thermodynamic limit. The Hamiltonian for N spins may be written

$$\mathscr{H} = -\frac{1}{2}\frac{J}{N}\left(\sum_{i=1}^{N} s_i^2\right) - \bar{H}\sum_{i=1}^{N} s_i \tag{4.2}$$

The thermodynamics of the model (and references to the literature) are given by Thompson.[20] Via a Kac–Hubbard transformation one sees that the partition function is proportional to

$$\mathscr{Z}_N(H,T) = \int_{-\infty}^{\infty} \exp\left[Ng(\mu;K,\bar{h})\right] d\mu \tag{4.3}$$

with $K = J/k_B T$, $\bar{h} = \bar{H}/k_B T$, and

$$g(\mu) = -\tfrac{1}{2}K\mu^2 + \ln\cosh(K\mu + \bar{h}) \tag{4.4}$$

For $T < T_c$ (given by $K_c = 1$) and $\bar{H} \geqslant 0+$ the magnetization, $m(\bar{H}, T)$, is the positive root of

$$m = \tanh(Km + \bar{h}) \tag{4.5}$$

while for $\bar{H} \leqslant 0-$ the negative root is to be chosen. For large N one easily sees that the transition is rounded on the scale $\bar{H}_N \sim 1/N$ which corresponds to the previous block limit since N fills the role of V even though the model has no proper spatial geometry. Indeed, for small $\bar{h}$ one finds the singular behavior of the free energy per spin is

$$\left[f(\bar{H},T) - f_\infty(0,T)\right]_s = -N^{-1}\ln\left[2\cosh(m\bar{h}N)\right] \tag{4.6}$$

with $f_\infty(\bar{H},T) = g[m(\bar{H},T); \bar{H}, T]$; this is precisely analogous to the previous "two-peak" scaling results, (2.18) and (3.14), in the block limit. The corrections to this leading behavior follow by using the method of steepest descents to evaluate (4.3) and are thus of relative order $1/N$, $1/N^2$, etc. By way of example one finds that the initial susceptibility is

$$\chi_0(T,N) = m_0^2(T)N/k_BT + \chi_\infty\left[1 - 2K + 2J(1-K)\chi_\infty\right] + O(N^{-1}) \tag{4.7}$$

where the limiting bulk susceptibility is $\chi_\infty = (k_BT\cosh^2 Km - J)^{-1}$.

Although the two-peak result is confirmed by the infinite range model and, more pertinently, by the transfer matrix analysis, it cannot be regarded as beyond question except, perhaps, in the low-temperature region, $\exp(-2dJ/k_BT) \ll 1$, where only few excitations are present. Indeed, the question of the probability distribution of the magnetization (especially in Heisenberg spin systems) is a matter of some subtlety.[25,27,30–32]

In the cylinder limit $L_\parallel/\xi_\parallel \gg 1$, the rounding is on the scale $H_A \simeq k_BT/m_0A\xi_\parallel$ so that the behavior of $\xi_\parallel(T; L_\perp)$ is of interest. In the low-temperature limit one easily shows that the prefactor in (3.4) behaves for all d as

$$D(T; L_\perp) \to \tfrac{1}{2}, \qquad \text{for} \quad T \to 0, \quad L_\perp \text{ fixed} \tag{4.8}$$

Since $m_0(T)$ and $\Sigma(T)$ also approach constants when $T \to 0$ the rounding thus varies as

$$H_A(T; L_\perp) \propto T\exp(-\Sigma_0 A/k_BT)/A \tag{4.9}$$

when $T \to 0$ at fixed $A = L_\perp^{d-1}$. This result is implicit also in Refs. 10 and 11; however, the arguments of Blöte, Nightingale, and Cardy[9,13] yield rounding on the scale $T\exp(-\Sigma_0 A/k_BT)/L_\perp^d$ which is smaller than (4.9) by a factor $1/L_\perp$. The reason for this discrepancy is explained in Appendix B: qualitatively, the inadequacy of their argument is the use of renormalization group flows *linearized* around the discontinuity fixed point at $T, H = 0$, $L_\perp = \infty$, to renormalize down to finite $L_\perp$.

The low-temperature limiting form (4.9) is valid only while the total probability, $A \exp[-2(d-1)J/k_B T]$, of a layer excitation is small: see region (i) of Fig. 3. In the opposite limit, namely, $L_\perp \to \infty$ at fixed $T > 0$ (with $T < T_c$), (4.8) fails. The analysis of Appendix A shows, for the two-dimensional Ising model, that one may then write

$$\frac{\xi_\parallel(T; L_\perp)}{\xi_\infty(T)} \approx P(T)\left(\frac{A}{\xi_\infty^{d-1}}\right)^\omega \exp\left[L_\perp^{d-1}\sigma(T)\right] \tag{4.10}$$

where $P(T)$ is a slowly varying function with a finite limit at $T = T_c-$, while $\omega(d=2) = \frac{1}{2}$. It is reasonable to conjecture that this expression with an appropriate exponent $\omega(d)$ or, possibly, different functional form, holds generally for $L_\perp \to \infty$ at T fixed for $d \leqslant d_> = 4$ (see Section 5 for $d > d_>$): see region (ii) in Fig. 3 which is bounded, following (3.15), also by $L_\perp \sim \xi_\infty(T)$. The rounding is then smaller than (4.9) by a factor $1/A^\omega$ and is also "enhanced" by a factor $\xi_\infty^{\omega(d-1)-1}$ which diverges as $T \to T_c-$ if $\omega(d-1) > 1$. Note, further, that if one invokes the hyperscaling relation,[33-35] $\mu = (d-1)\nu$, for the surface tension exponent defined via

$$\sigma(T) = \Sigma(T)/k_B T \approx B_\sigma\left[(T_c - T)/T_c\right]^\mu, \qquad \text{for} \quad T \to T_c- \tag{4.11}$$

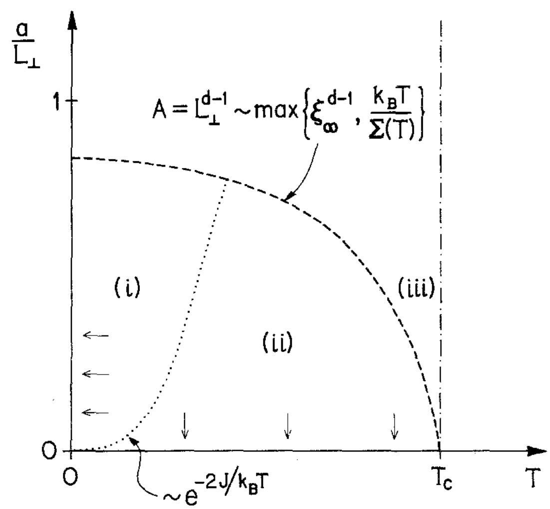

Fig. 3. Sketch of the region, (i) with (ii), of validity of asymptotic scaling at the first-order transition in the cylinder limit $L_\parallel \gg \xi_\parallel$ for cross-section area $A = L_\perp^{d-1}$. The arrows indicate the order of limits appropriate in the two distinct regimes (i) and (ii). The domain of critical behavior of $\xi_\parallel(T, L_\perp)$ is marked (iii).

the ratio $\xi_{\|}/\xi_{\infty}$ becomes, near criticality, a function only of the combination $L_{\perp}/\xi_{\infty}$. This accords with the natural scaling hypothesis for the critical region advanced in the following section. Hyperscaling is valid only for $d \leqslant 4$, although (3.4) should hold above $d = 4$: the modifications of naive finite-size scaling needed when hyperscaling fails are also discussed in the next section.

5. FIRST-ORDER SCALING NEAR CRITICALITY

We consider now how the finite-size scaling description of a first-order transition in the block and cylinder limits goes over, as it should, into standard finite-size scaling theory[1-3,16] for critical points when $T \to T_c -$. We also address the situation above $d = 4$ dimensions where naive finite-size scaling and hyperscaling breakdown[16] because, as we will demonstrate, a *dangerous irrelevant variable*[36] must be taken into account.

The finite-size scaling hypothesis for the critical region asserts that all unbounded lengths, say, L, should be scaled by the correlation length $\xi_{\infty}(T)$ which diverges as $|t|^{-\nu}$ with $t = (T - T_c)/T_c$. With the notation of Section 2, namely, $V = L_0^d$ and $l_j = L_j/L_0$, we may thus write the hypothesis

$$f_c(H, T; L_j) \approx |t|^{2-\alpha} W_c\big(h/|t|^{\Delta}, L_0|t|^{\nu}; l_j\big) \tag{5.1}$$

with $\Delta = \beta + \gamma$. Note that, as usual, analytic background terms have been subtracted from $f(H, T; L_j)$ to define f_c: for simplicity we have also dropped the metrical or amplitude factors A, B, etc. which permit one to normalize the scaling function, $W_c(y, z; l_j)$ [compare with (2.6)].

Now, if the neglect of irrelevant variables is permissible, (5.1) should be valid in the full domain $H, t, L_0^{-1} \to 0$. But this critical scaling domain overlaps the region of validity of the first-order scaling results (2.27), (3.14) and (3.19) for $f_s = f - f_{\infty}(H = 0, T)$. We recall that the first-order scaling involved the two combinations

$$y_V = m_0 h V \approx B h L_0^d |t|^{\beta}, \qquad \text{as} \quad T \to T_c - \tag{5.2}$$

where (2.7), (2.10) and (2.16) have been used, and

$$x = \frac{L_{\|}}{\xi_{\|}(L_{\perp})} \approx p_c \frac{L_{\|}|t|^{\nu-(d-1)\omega\nu}}{A^{\omega}} \exp\left[-\frac{\sigma(T) L_0^{d-1}}{l_1}\right] \tag{5.3}$$

where p_c is a constant and we have used (2.1) and (4.10). At first sight these scaling variables are not in the scaling form (5.1): however, if *hyperscaling* holds,[32-35] so that $d\nu = \Delta + \beta$ and $\mu = (d-1)\nu$, it is easy to check the relation

$$y_V \approx Byz^d \qquad \text{with} \quad y = h/|t|^{\Delta} \quad \text{and} \quad z = L_0|t|^{\nu} \tag{5.4}$$

as $T \to T_c -$ and, likewise, using (4.11),

$$x \approx X(z; l_1) = p_c l_1^{1+\omega} z^{1-(d-1)\omega} \exp(-B_\sigma z^{d-1}/l_1) \tag{5.5}$$

where we recall that $\omega = \frac{1}{2}$ for $d = 2$ (by Appendix A) but is not determined for $d > 2$. (One might, indeed, have $z^{(d-1)\omega}$ replaced by a different, but still slowly varying function of z for $d = 3$.)

Having checked that, indeed, the first-order scaling forms for $T < T_c$ do respect critical scaling when $d < d_{>} = 4$ we can use (3.19) to conclude that the critical scaling function, $W_c(y, z; l_j)$, behaves in accord with

$$W_c(y, z; l_j) - W_c(0, z; l_j) \approx -\ln 2 \cosh\left[B^2 y^2 z^{2d} + \tfrac{1}{4} X^2(z; l_1)\right]^{1/2} \tag{5.6}$$

as $z \to \infty$ (with $T < T_c$); this expression encompasses both the block and cylinder limits and also the crossover between them i.e., the full range $L_{\|} \equiv L_1 \gtrsim L_j$ for $j = 2, 3, \ldots$. It is remarkable that such an explicit result can be found for the finite-size scaling function in the critical region!

An interesting point raised by Brézin[16] for the cylinder limit is relevant to phenomenological renormalization group calculations[4]; this concerns the finite-size scaling properties of the spectral gap $\Delta f(0, T; L_\perp)$ or, equivalently, of $\xi_{\|}(T; L_\perp)$ in the critical region and, in particular, at the critical point. By the general finite-size scaling hypothesis one anticipates

$$\xi_{\|}(H, T; L_\perp) \approx |t|^{-\nu} Z_{\|}(h/|t|^{\Delta}, L_\perp |t|^{\nu}) \tag{5.7}$$

where we have included a dependence on the field but, for simplicity, supposed $L_2 = \cdots = L_d = L_\perp$ (while $L_{\|} \to \infty$). A check on this scaling ansatz is also provided by (4.10) which, if hyperscaling holds, yields

$$Z_{\|}(0, z) \approx p_{\|} z^{(d-1)\omega} \exp(B_\sigma z^{d-1}) \tag{5.8}$$

as $z \to \infty$, where $p_{\|}$ is a constant. At the critical point itself, (5.7) then gives

$$\xi_{\|,c}(L_\perp) \equiv \xi_{\|}(0, T_c; L_\perp) \sim L_\perp = A^{1/(d-1)} \tag{5.9}$$

which, of course, also follows directly from the finite-size scaling principle. However, if hyperscaling fails, as it does for $d > d_{>} = 4$, the general relation (3.4) for $\xi_{\|}$ is *inconsistent* with the scaling ansatz (5.7) which must, thus, be restricted to $d < 4$. This conclusion is, indeed, confirmed by Brézin's calculations[16] for the multicomponent limit $n \to \infty$. (Note that for finite systems, especially with nonperiodic boundary conditions, this limit must be distinguished from the standard spherical model.[1,37])

One may also define an *overall* size-dependent correlation length,[16] say, $\xi_V(H, T; L_j)$, in the block limit via the general scaling relation

$$\xi_V/a = \left(k_B T\chi/a^d\right)^{\nu/\gamma} \tag{5.10}$$

where $\chi \equiv \chi(H, T; L_j)$ is the finite-size susceptibility. The thermodynamic scaling relation (5.1) then leads to the analog of (5.7), namely,

$$\xi_V(H, T, L_j) \approx |t|^{-\nu} Z_V\left(h/|t|^\Delta, L_0|t|^\nu; l_j\right) \tag{5.11}$$

and hence, formally, as $y = h/|t|^\Delta$ and $z = L_0|t|^\nu \to 0$, to

$$\xi_{V,c}(L_j) \equiv \xi_V(0, T_c; L_j) \sim L_0 = V^{1/d} \tag{5.12}$$

which parallels (5.9). [If one defines ξ_V via $(f_c)^{-1/d}$, which is dimensionally appropriate, one would obtain (5.11) and, thence, (5.12) only when the hyperscaling relation $2 - \alpha = d\nu$ holds.]

Now Brézin concluded[16] by direct calculation for $n \to \infty$ that both (5.9) and (5.12) fail for $d > 4$, the length exponents no longer being unity but, rather, depending on d. Since mean field theory holds for the bulk system when $d > d_>$, one might hope to obtain the correct results by appropriate scaling arguments without the need for explicit calculation. To that end, we appeal to our observation that in the first-order cylinder limit ($L_\| \to \infty$) the finite dimensions enter principally through the fluctuation energy necessary to create an interface across the system and hence through the dimensionless combination

$$y_\Sigma = \frac{A\Sigma(T)}{k_B T} \approx B_\sigma L_\perp^{d-1}|t|^\mu = B_\sigma\left(L_\perp|t|^{3/2(d-1)}\right)^{d-1} \tag{5.13}$$

in which, for $d > 4$, we have used the mean-field result $\mu = 3/2$. This suggests that for $d > 4$, where $\nu = 1/2$, and $\Delta = 3/2$, the scaling ansatz (5.7) should be replaced by

$$\xi_\|(H, T; L_\perp) \approx |t|^{-1/2} Z_\|\left(h/|t|^{3/2}, L_\perp|t|^{3/2(d-1)}\right) \tag{5.14}$$

At criticality this yields

$$\xi_{\|,c}(L_\perp) \sim a(L_\perp/a)^{(d-1)/3} \sim A^{1/3} \tag{5.15}$$

providing the scaling function is well behaved. This surmise is, in fact, confirmed precisely by Brézin's $n \to \infty$ calculations![16] [On the borderline $d = 4$ Brézin obtains $\xi_{\|,c} \sim L_\perp[\ln(L_\perp/a)]^{1/3}$.]

The parallel argument for the block limit suggests that only the bulk combination

$$y_V = m_0 hV \quad \text{scaling as} \quad t^{\beta+\Delta} L_0^d = \left(L_0 t^{2/d}\right)^d \tag{5.16}$$

with $\beta + \Delta = 2$ for $d > 4$, should be important so that for $d > 4$, one should replace (5.11) by

$$\xi_V(H, T; L_j) \approx |t|^{-1/2} Z_V\big(h/|t|^{3/2}, L_0|t|^{2/d}; l_j\big) \tag{5.17}$$

where the l_j are fixed. At criticality this leads to

$$\xi_{V,c}(L_j) \sim a(L_0/a)^{d/4} \sim V^{1/4} \tag{5.18}$$

which also agrees precisely with Brézin's result for $n \to \infty$.

From these considerations we learn that correlation lengths all scale in the usual way with $|t|^{-\nu}$ but that the overall dimensions of a system enter, in the block limit, through the total bulk ordering free energy ratio $y_V = F_V/k_B T$ and, in the cylinder limit, through the interfacial free energy ratio $y_\Sigma = F_\Sigma/k_B T$. Below the borderline $d_> = 4$ hyperscaling prevails and *all* lengths and distances, including the correlation lengths ξ_∞, $\xi_\|$, etc. scale in the same way with $t^{-\nu}$; above $d = 4$, however, the dimensions L_j scale with new powers of t.

It is instructive to place these results within a renormalization group context[38,39] in which *all* lengths and distances should renormalize with the spatial rescaling factor b as

$$f_c(h, t, u; R) \approx b^{-d} f_c(b^{\lambda_h} h, b^{\lambda_t} t, b^{\lambda_u} u; R/b) \tag{5.19}$$

Here R denotes collectively the various lengths and distances and we have included the parameter u which derives from the strength of the fourth-order term in the Landau–Ginzburg–Wilson effective Hamiltonian

$$\mathscr{H}[s]/k_B T = \int d^d x \left[\tfrac{1}{2}(\nabla s)^2 - h_0 s + \tfrac{1}{2} r_0 s^2 + u_0 s^4 \right] \tag{5.20}$$

with $r_0 \sim t$. Further we have chosen to renormalize in the standard way which keeps the coefficient of $(\nabla s)^2$ constant.

For $d < 4$ one anticipates a nontrivial fixed point[38,39] and u in (5.19) represents the deviation from the fixed point value, $u^* > 0$, the corresponding eigenvalue, λ_u being negative. The choice $b = |t|^{-1/\lambda_t}$, as $t \to 0$, leads to the standard scaling form where f_c scales as $|t|^{d/\lambda_t}$, h scales as $|t|^{\lambda_h/\lambda_t}$, and R scales as $|t|^{-1/\lambda_t}$. If we introduce the correction-to-scaling exponent $\theta = -\lambda_u/\lambda_t > 0$ we see that the field u enters only in the combination $u|t|^\theta$ which *vanishes* as $t \to 0$, so that u is an *irrelevant variable*. One thus obtains formally the standard exponent identifications, $2 - \alpha = d\nu = d/\lambda_t$ and $\Delta = \lambda_h/\lambda_t$, which entail hyperscaling *and* the scaling of all lengths as $|t|^{-\nu}$. The main tacit assumption is that u is a *harmless* irrelevant variable which may be set equal to zero (corresponding to u_0 at the fixed point value u^*) without resulting in any singular or anomalous behavior.

When the dimensionality exceeds $d_> = 4$ the critical behavior is controlled by the Gaussian fixed point[38,39] and one has the renormalization group eigenvalues

$$\lambda_h = \tfrac{1}{2}d + 1, \quad \lambda_t = 2, \quad \text{and} \quad \lambda_u = 4 - d \tag{5.21}$$

The last of these is still negative so that u is again an irrelevant variable. However, one now has $u^* = 0$ so that coefficient $u \cong u_0$ cannot be allowed to vanish for obvious reasons of stability (at and below criticality). If one modifies the renormalization group transformation so that the spin is rescaled (with b) so as to keep u fixed at u_0 one finds that the coefficient of the $(\nabla s)^2$ term now rescales rapidly to $+\infty$ as b grows. This justifies the classical saddle point approximation, equivalent to mean field theory, which, for a homogeneous bulk system, yields

$$f_c \approx \min_s \left\{ -h_0 s + \tfrac{1}{2} r_0 s^2 + u s^4 \right\} \tag{5.22}$$

Rescaling by putting $s = \tilde{s}/u^{1/2}$ shows that u enters the free energy in the form

$$f_c(h, t, u) = u^{-1} \tilde{f}(h u^{1/2}, t) \tag{5.23}$$

and, thus the magnetization, m, has a prefactor $u^{-1/2}$. The divergence of f and m as $u \to 0$ confirms the *dangerous* character[36] of the variable u. One may again use the general renormalization group relation (5.19) with the choice $b = |t|^{-1/\lambda_t} = |t|^{-1/2}$ but must recognize that f_c and h will entail factors of u as in (5.23). One thus anticipates

$$f_c \approx \frac{|t|^{d/2}}{u|t|^\theta} W_0 \left[\frac{h(u|t|^\theta)^{1/2}}{|t|^{(d+2)/4}}, u|t|^\theta; R|t|^{1/2} \right] \tag{5.24}$$

where the correction-to-scaling exponent is now $\theta = \frac{1}{2}(d-4) > 0$. If u is fixed at its initial physical value, u_0, it is readily checked that (5.24) represents the standard, classical scaling prediction with $\alpha = 0$, $\Delta = 3/2$ and $\nu = 1/2$. From this one likewise finds the bulk, zero-field behavior,

$$m_0 \approx B_0 |t|^{1/2}/u^{1/2} \quad \text{and} \quad \chi = C_0^{\pm}/|t| \qquad (t \gtrless 0) \tag{5.25}$$

where all the u dependence has been displayed; however, for *finite systems* the variable u may still enter in combination with the various dimensions, as we now indicate.

The bulk or block ratio can be written

$$y_V = f_c V \approx W_{00} \frac{|t|^2}{u} L_0^d = W_{00} \frac{(L_0|t|^{1/2})^d}{(u|t|^\theta)} = W_{00} \frac{(L_0|t|^{2/d})^d}{u} \tag{5.26}$$

where $W_{00} = W_0(0, 0)$ and the dangerous irrelevant scaling combination $u|t|^\theta$ has been explicitly isolated. This likewise exhibits length scaling as $|t|^{-\nu}$ but shows that L_0 scales with the "anomalous" power $|t|^{-2/d}$ when

$u \equiv u_0$ is fixed, as proposed in (5.16). By the same token $u|t|^\theta$ must now enter (5.11) and (5.24) in a singular fashion i.e., as a divisor of $L_0|t|^\nu$. For the corresponding interfacial or cylinder ratio note, first, that $\Sigma(T)$ is quite generally,[33,34] proportional to $m_0^2\xi_\infty/\chi$ when $t \to 0$. Thus we have

$$y_\Sigma = \frac{\Sigma A}{k_B T} \approx \frac{B_0^2|t|}{u} \frac{|t|}{C_0^-} \frac{a_0 L_\perp^{d-1}}{|t|^{1/2}} = D_0 \frac{\left(L_\perp |t|^{1/2}\right)^{d-1}}{\left(u|t|^\theta\right)}$$

$$= \frac{D_0}{u}\left(L_\perp |t|^{3/2(d-1)}\right)^{d-1} \tag{5.27}$$

where $D_0 = a_0 B_0^2 / C_0^-$ and a_0 is a constant. Again we see standard renormalization group scaling with the dangerous irrelevant combination, $u|t|^\theta$, leading to the anomalous cross-sectional exponent $3/2(d-1)$, as advanced in (5.14).

In summary, our discussion of finite size scaling for $d > d_> = 4$ suggests that the *full* scaling form, encompassing single-phase behavior and the block and cylinder first-order transition limits will entail the *three* scaled variables

$$L_j |t|^{1/2}, \quad L_0^d t^2 \equiv V t^2, \quad \text{and} \quad L_\perp^{d-1} |t|^{3/2} \equiv A|t|^{3/2} \tag{5.28}$$

Thus all, for example, should enter as arguments in (5.7), (5.11), (5.14), (5.17), and (5.24) for a complete asymptotic description. By contrast, when $d < 4$ all limits are covered by the single combination $L_j|t|^\nu$.

6. PHENOMENOLOGICAL RENORMALIZATION USING SINGLE-PHASE FUNCTIONS

The phenomenological renormalization technique is a numerical method for studying bulk criticality which has been applied successfully to several two-dimensional problems. (See Ref. 4 for a recent review.) Various versions of the method have been used since the approach was originally introduced by Nightingale.[40] For simplicity we consider here only $d = 2$ and the $L \times \infty$ geometry and recall one of the most frequently used techniques. Thus finite-size scaling for, say, the zero-field susceptibility implies that the relation

$$\frac{\ln[\chi_L(T_0)/\chi_{L'}(T_0)]}{\ln[\chi_{L''}(T_0)/\chi_{L'}(T_0)]} \approx \frac{\ln(L/L')}{\ln(L''/L')} \tag{6.1}$$

is valid asymptotically as L, L' and $L'' \to \infty$ when $T_0 = T_c$. Conversely, the solution, $T_0(L, L', L'')$, of this relation regarded as an equality approximates the true critical point, T_c. Given a reliable estimate for T_c one may likewise estimate the exponent ratio γ/ν from

$$\gamma/\nu \simeq \ln[\chi_L(T_c)/\chi_{L'}(T_c)]/\ln(L/L') \tag{6.2}$$

Similar calculations have been performed using the specific heat and using $\xi_{\parallel}(T)$; in the latter case ν may be estimated by linearizing around the phenomenological fixed point.[40] More recently, Hamer[41] has suggested a variant using a function which approaches the spontaneous magnetization.

Now the "single-phase" free energy functions,[10,11] $f_{\pm}(H,T;L)$, defined in (3.11) should obey finite-size scaling because they are obtained by algebraic manipulation of quantities, namely, $f_0(H,T;L)$ and $f_1(H,T;L)$, which should also obey scaling. This entails the new assumption that the first excited free energy level, $f_1(H,T;L)$, obeys finite-size scaling; however, this is certainly plausible since f_1 and f_0 are branches of the same analytic function[27] of H. Notice that the single-phase free energies $f_{\pm}(H)$, are not even functions of H, unlike $f_0(H)$. Thus the corresponding scaling functions for *all* the field derivatives of $f_{\pm}(H)$ have nonvanishing values at $H=0$. Therefore the $f_{\pm}(H)$ can be used in phenomenological renormalization calculations designed to study *odd* derivatives of the free energy *at* the first-order boundary when $T \to T_c -$ (as well as higher-order even derivatives).

To illustrate the potentialities of this observation, we report here numerical calculations for the nearest-neighbor square and triangular lattice Ising models (with isotropic interactions). We use periodic boundary conditions and consider the expansion

$$m_+^{(L)}(T_c, H) = -\left(\frac{\partial f_+^{(L)}}{\partial h}\right)_c = \sum_{k=0}^{\infty} c_k^{(L)} H^k \tag{6.3}$$

From the scaling ansatz we have, for $T \simeq T_c -$ and small H,

$$m_+^{(L)}(T,H) \approx D_1 L^{-\beta/\nu} Y_+(D_2 t L^{1/\nu}, D_3 H L^{\Delta/\nu}) \tag{6.4}$$

where the scaling function $Y_+(x,y)$ should be universal and, hence, independent of lattice structure with a proper assignment of the nonuniversal amplitudes, or metrical factors, D_i. Then the critical point expansion coefficients should vary as

$$c_k^{(L)} \approx D_1 D_3^k L^{(k\Delta-\beta)/\nu} (\partial^k Y_+/\partial y^k)_0 / k! \tag{6.5}$$

when $L \to \infty$, where the subscript 0 denotes evaluation at $x=y=0$. With this observation in mind we have calculated numerically the approximants

$$(\gamma_k/\nu)^{(L)} = \ln[c_k^{(L)}/c_k^{(L-a)}]/\ln[L/(L-a)] \tag{6.6}$$

which, as $L \to \infty$, should converge to γ_k/ν, where

$$\gamma_k = k\Delta - \beta = \gamma + (k-1)\Delta \tag{6.7}$$

so that $\gamma_0 = -\beta$, $\gamma_1 = \gamma$, $\gamma_2 = \gamma + \Delta = \beta + 2\gamma$, etc.

For planar Ising models we have,[14,42,43] $\beta = \frac{1}{8}$, $\gamma = 1\frac{3}{4}$, and $\nu = 1$ and the validity of the relations for γ_k with $k = 2, 3, \ldots$ is well established in series extrapolation studies.[44–46] As a test of the approach, therefore, we list in Table I data for $(\gamma_k/\nu)^{(L)}$ for $L/a = 8$, 9 and 10 and several values of k, for the square and the triangular lattices. The calculations yielding these data are standard and will not be discussed here.[4,11] The expansion coefficients $c_k^{(L)}$ are obtained by numerical differentiation of $f_+(H, T_c; L)$ which, as a result of roundoff errors, restricts k to 8 for the triangular lattice and 6 for the square lattice. The values of $(\gamma_k/\nu)^{(L)}$ have been truncated to a sufficient number of places to display the nature of the convergence to γ_k/ν, and are otherwise accurate to the order displayed or higher. (Only in the last few entries for $L = 10a$ do roundoff errors manifest themselves.)

Frequently such phenomenological renormalization data for critical exponents are fitted well by the form $|(\gamma_k/\nu)^{(L)} - \gamma_k/\nu| \propto L^{-p_k}$ with an effective power law convergence exponent p_k of magnitude[4,47] around 2. Here, however, an analysis of $(\gamma_k/\nu)^{(L)}$ for $L/a = 6, 7, \ldots, 10$ does not reveal any regular pattern of convergence: indeed the $k = 0$ sequence, $(\beta/\nu)^{(L)}$, for the triangular lattice is not even monotonic! [Note that the restriction to $L/a \leqslant 10$ is the price paid for studying high-order derivatives so that $f_+(H)$ must be computed to high accuracy; when only low-order derivatives are of interest it is feasible[4] to go up to $L/a \lesssim 16$.] Extrapolation of exponent data using special methods[9,48] to accelerate convergence may in some cases[4] improve agreement with conjectured or exact values (if known) by an order of magnitude. However, we have not attempted to perform any such extrapolations because the data here are clearly not in a regime of asymptotic convergence.[47] Instead, let us discuss further those features which may be of interest in applications to other models.

First, note that the accuracy of critical exponents estimates for the higher-order derivatives is no worse (perhaps even somewhat better) than for the second derivative. This is not necessarily surprising: the higher-order derivatives of the free energy are more difficult to calculate accurately but the rate of their convergence to the $L \to \infty$ limit, which is determined by corrections to the leading finite-size scaling behavior, should not be qualitatively different. However, the values of $(\beta/\nu)^{(L)}$ are relatively closer to β/ν than are the other exponents estimates. (See also Hamer.[41]) There is also a striking difference in the accuracy of the estimates found for the two different lattices, which is probably related to the fact that the triangular lattice is more closely packed.

Table I. Values of $(\gamma_k/\nu)^{(L)} - \gamma_k/\nu$ for the Triangular and the Square Lattices[a]

	Triangular				Square			
k	$L = 8a$	$L = 9a$	$L = 10a$		$L = 8a$	$L = 9a$	$L = 10a$	
0	− 2.02	− 2.12	− 1.94	$\times 10^{-6}$	6.10	4.62	3.63	$\times 10^{-4}$
1	4.56	3.63	2.97	10^{-2}	8.16	6.61	5.48	10^{-2}
2	− 2.64	− 1.93	− 1.46	10^{-3}	− 2.42	− 2.00	− 1.68	10^{-2}
3	− 8.95	− 7.98	− 6.74	10^{-4}	− 7.39	− 5.91	− 4.82	10^{-2}
4	13.6	4.99	1.30	10^{-4}	− 13.6	− 10.9	− 8.83(± 1)	10^{-2}
5	5.62	2.98	1.69	10^{-3}	− 2.27	− 1.82	− 1.48	10^{-1}
6	18.2	10.4	6.37	10^{-3}	− 4.13	− 3.33	− 2.8(± 1)	10^{-1}
7	− 4.82	− 3.01	− 1.97(± 1)	10^{-1}				
8	− 21.8	− 13.4	− 8.6(± 1)	10^{-3}				

[a] The values are rounded to three figures or roundoff errors are indicated. The powers of 10 represent constant factors for each entry in the row.

We may also use our data to estimate the scaling function $Y_+(x, y)$ in (6.4). There is some arbitrariness in fixing the metrical factors D_i but a convenient approach[49] is to specify them by requiring

$$Y_+(x, y) = 1 + x + y + O(xy, x^2, y^2) \tag{6.8}$$

Then the expansion coefficients, Y_k, in

$$Y_+(0, y) = 1 + y + Y_2 y^2 + Y_3 y^3 + \cdots \tag{6.9}$$

should be universal and are approximated by

$$Y_k^{(L)} = \frac{c_k^{(L)}}{c_0^{(L)}} \left(\frac{c_0^{(L)}}{c_1^{(L)}} \right)^k \tag{6.10}$$

Our numerical results for both the square and triangular lattices are summarized by

$$\begin{gathered} Y_2 = -8.35 \pm 0.30, \qquad Y_3 = 104 \pm 7 \\ Y_4 = -(1.53 \pm 0.16) \times 10^3, \qquad Y_5 = (2.35 \pm 0.35) \times 10^4, \\ Y_6 = -(3.0 \pm 0.7) \times 10^5 \end{gathered} \tag{6.11}$$

The coefficients alternate in sign regularly to this order but some preliminary calculations suggest that this does not continue in higher orders.

As regards verifying the expected universality of $Y_+(0, y)$, the coefficients Y_k do not provide an optimal test. The problem is that parameters calculated from the second free-energy derivative data ($k = 1$) seem, as mentioned, to converge more slowly than for other derivatives. Owing to the factor $(c_1^{(L)})^k$ in (6.10), the associated uncertainties are amplified in the estimation of the Y_k. As an alternative test of universality it is better to study the ratios

$$R_k^{(L)} = c_k^{(L)} c_{k-2}^{(L)} / \left(c_{k-1}^{(L)}\right)^2 \tag{6.12}$$

which involve only low powers of the $c_j^{(L)}$, thus avoiding error accumulation. These ratios serve to approximate the universal scaling function coefficient ratios

$$R_k = Y_k Y_{k-2} / (Y_{k-1})^2, \qquad k = 2, 3, \ldots \quad \text{with} \quad Y_0 = Y_1 = 1 \tag{6.13}$$

In Table II we list several $R_k^{(L)}$ for $L/a = 9$ and 10. It appears again that the $R_k^{(L)}$ for $L/a = 6, 7, \ldots, 10$ do not yet exhibit any simple pattern of convergence although the triangular lattice values appear to be more rapidly convergent. The values for $L = 10a$ for the two lattices agree to within 4% for $k = 2$ and 6 but to better than 1% for $k = 3$, 4, and 5: this represents a gratifying confirmation of the anticipated universality.

Table II. The Ratios $R_k^{(L)}$ for the Triangular and Square Ising Lattices[a]

	Triangular		Square	
k	$L = 9a$	$L = 10a$	$L = 9a$	$L = 10a$
2	−4.20	−4.17	−4.39	−4.33
3	0.988	0.991	0.979	0.983
4	0.893823	0.893825	0.898477	0.897653
5	0.82861	0.82868	0.83817	0.83647
6	0.6922	0.6925	0.7211	0.7159

[a] The values have been rounded to the last place displayed.

As a final comment on numerical methods using the cylinder geometry, note that the position of a first-order phase boundary might be located[8,9] by searching for points at which the susceptibility (or the specific heat) diverges exponentially with the cross-sectional area $A = L_\perp^{d-1}$. Similarly, one could examine the spectral gap $1/\xi_\parallel(H, T; L) = A(f_1 - f_0)$ as a function of H. By the analysis of Section 3 this has the scaling form

$$1/\xi_\parallel^2(H, T; L_\perp) \approx (2Am_0h)^2 + 1/\xi_\parallel^2(0, T; L_\perp) \tag{6.14}$$

for small $h = H/k_BT$. If one uses the standard phenomenological rescaling approach (following, e.g., Rikvold *et al.*[50]) and solves the relation

$$\xi_\parallel(H, T; L_\perp)/L_\perp = \xi_\parallel(H, T; L'_\perp)/L'_\perp \tag{6.15}$$

for H one then finds that the phase boundary is given correctly up to errors of order $1/\xi_\parallel(L_\perp)$ and $1/\xi_\parallel(L'_\perp)$, which are exponentially small in $L_\perp$ and $L'_\perp$. (At a continuous transition one expects errors decaying asymptotically[47] as $L_\perp^{-(1+\theta)/\nu}$ where θ is the leading singular correction-to-scaling exponent.) One can see similarly that the effective renormalization group eigenvalue, $\lambda^{(L)} = 1/\nu^{(L)}$, as computed formally by the usual rule,[4] diverges like $A = L_\perp^{d-1}$ at a first-order transition. This sort of behavior has, indeed, been observed[9] in a study of the first-order thermal transition in planar q-state Potts models with $q > 4$.

7. CONCLUDING REMARKS

Finite-size effects at a first-order phase boundary evidently provide a rich panorama of phenomena, involving a profound dependence on the shape of a system. We have focused mainly on just two geometries, namely, block and cylinder, and have studied the crossover between them. However, there are other geometries of significant theoretical and experimental interest, for example, the slab geometry where finite-size and, especially, surface phenomena have been a subject of intensive but by no means exhaustive study.[51]

It should also be emphasized that we have considered only boundary conditions which do not break the $H \Leftrightarrow -H$ symmetry. Since the shift in the first-order transition due to symmetry breaking surface effects is asymptotically larger than the rounding,[5] further interesting effects may arise with more realistic, free or pinned boundary conditions.[28,29] In particular the surface contributions themselves become important.[1,28,29]

Finally, recall that we have discussed only scalar (and discrete) spin systems. For systems with vector (and/or continuous) spins less information on the rounding of the first-order transitions is available but important qualitative differences arise[13,16] from the different dependence of $\xi_\parallel$ on $L_\perp$ which reflects the replacement of relatively sharp domain walls by indefinitely diffuse Bloch walls.

ACKNOWLEDGMENTS

We have enjoyed useful discussions with P. Nightingale and wish to thank M. Barma, E. Brézin, D. A. Huse, R. Pandit, and S. Trugman for instructive interactions. The financial support of the Rothschild Fellowship Foundation (for V.P.) is appreciated. The researches reported have also received support from the National Science Foundation (under Grant No. DMR-81-17011) and, in part, through the Materials Science Center at Cornell University.

APPENDIX A: SPECTRAL GAP FOR THE SQUARE LATTICE ISING MODEL

We outline the calculation of the asymptotic form of the spectral gap, $a/\xi_\parallel(T)$, for the square lattice Ising model with couplings $J_1 \equiv J_\parallel$ along the cylinder axis and J_2 between spins in the same layer. In terms of the variables $K_i = J_i/k_B T$ and K_i^* related by

$$\tanh K_i^* = \exp(-2K_i), \qquad i = 1,2 \tag{A1}$$

one has[14] $K_1 = K_2^*$ at $T = T_c$ and $K_1 > K_2^*$ for $T < T_c$. The surface tension for an interface normal to the axis is given by[14,15]

$$\bar{\sigma} \equiv a\sigma(T) = a\Sigma(T)/k_B T = 2(K_1 - K_2^*) \tag{A2}$$

while the spectral gap for a lattice of width $M = L_\perp/a$ sites is[14]

$$\frac{a}{\xi_\parallel} \equiv \ln\frac{\Lambda_0}{\Lambda_1} = \frac{1}{2}\sum_{k=0}^{M-1}{}' \left[\gamma\left(\frac{2k+1}{M}\right) - \gamma\left(\frac{2k}{M}\right)\right] \tag{A3}$$

where for $T < T_c$ the function $\gamma(x)$ is analytic on the real axis, with period 2, being given by

$$\gamma(x) = \cosh^{-1}(\operatorname{ch} 2K_1^* \operatorname{ch} 2K_2 - \operatorname{sh} 2K_1^* \operatorname{sh} 2K_2 \cos \pi x) > 0 \tag{A4}$$

in which we have used the convenient abbreviations

$$\operatorname{ch} z \equiv \cosh z \quad \text{and} \quad \operatorname{sh} z \equiv \sinh z \tag{A.5}$$

Then if $c_r(T)$ is the rth Fourier coefficient of $\gamma(x; T)$ one has

$$a/\xi_{\parallel}(T, L_{\perp}) = -M \sum_{j=0}^{\infty} c_{(2j+1)M}(T) \tag{A6}$$

Because $\gamma(x)$ is analytic on the real axis the Fourier coefficients decrease exponentially with r and thus we have

$$a/\xi_{\parallel} \approx -Mc_M = -M \int_0^2 \gamma(x) \cos(M\pi x)\, dx \tag{A7}$$

where the corrections are of relative order c_{3M}/c_M which is of magnitude $\exp(-2\sigma L_{\perp})$ as may be seen from the explicit result below.

After an integration by parts, some algebra, and a shift of the contour of integration into the complex plane one obtains

$$a/\xi_{\parallel} \approx \pi^{-1}\left(e^{\bar\sigma} I_{M-1} - e^{-\bar\sigma} I_{M+1}\right) e^{-\bar\sigma M} \tag{A8}$$

up to exponential corrections, where

$$I_M = \int_0^{4K_2^*} \frac{e^{-M\tau}\, d\tau}{\left[\operatorname{ch}(\bar\sigma + \tau) - \operatorname{ch} \bar\sigma\right]^{1/2} \left[\operatorname{ch} 2(K_1 + K_2^*) - \operatorname{ch}(\bar\sigma + \tau)\right]^{1/2}} \tag{A9}$$

When $M = L_{\perp}/a$ is fixed, the limiting behavior of this integral as $T \to 0$ is obtained when $e^{-M\tau}$ varies slowly over $[0, 4K_2^*]$, that is, for $K_2^* \ll a/L_{\perp}$. One then finds that I_M approaches $\pi/\operatorname{sh} 2K_1$, which leads to

$$\xi_{\parallel}(T, L_{\perp}) \approx \tfrac{1}{2} a e^{2K_1 L_{\perp}/a} \tag{A10}$$

for $T \to 0$ with $L_{\perp}$ fixed.

By contrast, the behavior for large $L_{\perp}$ at fixed T in $(0, T_c)$ is obtained when $e^{-M\tau}$ varies rapidly or $L_{\perp}/a \gg 1/K_2^*$. In this limit we find

$$I_M \approx \left[\pi \operatorname{sh} 2K_1^* \operatorname{sh} 2K_2 / 2M \operatorname{sh} \bar\sigma\right]^{1/2} \tag{A11}$$

which finally leads to

$$\xi_{\parallel}(T; L_{\perp}) \approx \left(\frac{\pi}{2 \operatorname{sh} 2K_1^* \operatorname{sh} 2K_2}\right)^{1/2} \left(\frac{L_{\perp}/a}{\operatorname{sh} \bar\sigma(T)}\right)^{1/2} e^{L_{\perp}\sigma}\left[1 + O\left(\frac{1}{L_{\perp}}\right)\right] \tag{A12}$$

which is valid as $L_\perp/a \to \infty$ at fixed nonzero $T < T_c$. The crossover between the two limiting forms (A10) and (A12) occurs for $L_\perp K_2^*/a$ of order unity: see the dotted curve in Fig. 3 which thus corresponds to $L_\perp/a \approx \exp(2J_2/k_B T)$. The exponential corrections in (A7) lead to crossover to critical behavior as $A\Sigma(T)/k_B T$ becomes small where, in this case, $d = 2$ and so $A \equiv L_\perp$: see the dashed curve in Fig. 3. Via hyperscaling[35] this condition is equivalent to $L_\perp/\xi_\infty(T)$ of order unity as implied by the scaling criterion (3.15). To discuss the behavior of $\xi_\parallel(T, L_\perp)$ in the critical region, marked (iii) in Fig. 3, a more elaborate analysis is required which is not considered here.

APPENDIX B: RENORMALIZATION DOWN TO ONE DIMENSION

Blöte, Nightingale, and Cardy (BNC)[9,13] have discussed the effects of finite-size on the first-order transition with the aid of a low-temperature renormalization group rescaling approach. Here we present a brief critique of their arguments which demonstrates why they obtain results which, we believe, are not fully correct.

Let us first reformulate the BNC technique in a form suitable for comparing with our analysis in Sections 2–4. The basic idea of BNC is to use for finite-size systems the accepted *bulk* renormalization group recursion relations, as linearized about the $T, H = 0$ discontinuity fixed point, namely,[6,52]

$$h' = b^d h \quad \text{and} \quad T' = b^{1-d} T, \quad \text{so} \quad H' = bH \tag{B1}$$

where b is the standard spatial rescaling factor. For an $L_\parallel \times L_\perp^{d-1}$ geometry with $L_\parallel \geqslant L_\perp$ and periodic boundary conditions, these bulk relations are then supplemented with

$$L'_\parallel = L_\parallel / b \quad \text{and} \quad L'_\perp = L_\perp / b \tag{B2}$$

and the usual bulk flow equation for the free energy is extended by postulating

$$f_s(H, T; L_\parallel, L_\perp) \approx b^{-d} f_s(H', T'; L'_\parallel, L'_\perp) \tag{B3}$$

This framework is, of course, consistent with the standard asymptotic finite-size scaling hypothesis[1–3,5] and should thus be valid in the vicinity of the bulk first-order transition for *block* geometry. However, (B3) neglects, in particular, corrections due to nonlinearities of the renormalization group away from the fixed point (which can, at least in leading orders, be embodied in nonlinear scaling fields) as well as singular corrections to scaling. Nevertheless BNC *assume* that one may validly neglect all these corrections even as $L_\perp$ is renormalized down to $L'_\perp = a$, which corresponds to a *one*-dimensional chain. Thus they choose $b = L_\perp/a$ and, using (B3), assert

$$f_s(H,T;L_\parallel,L_\perp) \approx \left(\frac{a}{L_\perp}\right)^d f_s\left[H\frac{L_\perp}{a},T\left(\frac{a}{L_\perp}\right)^{d-1};\frac{L_\parallel}{L_\perp}a,a\right] \tag{B4}$$

where, it is argued, the right-hand side may be evaluated in terms of the free energy, $f_a(H',T';L')$, of a one-dimensional Ising model of length L'.

Now the two transfer matrix eigenvalues for a simple Ising chain with $\bar{h}' = H'a^d/k_BT$ and $K' = J/k_BT'$ are

$$\Lambda_0, \Lambda_1 = e^{K'}\left[\cosh\bar{h}' \pm (\sinh^2\bar{h}' + e^{-4K'})^{1/2}\right] \tag{B5}$$

In leading order for small T' and H' (effective control over higher orders having, in any case, been lost) the singular part of the free energy is thus given by

$$f_a \approx -\frac{1}{L'}\ln 2\cosh\left[(L'/a)(\bar{h}'^2 + e^{-4K'})^{1/2}\right] \tag{B6}$$

On appealing to (B4) with $L' \equiv L'_\parallel = L_\parallel a/L_\perp$, etc. one obtains a result which may be written as

$$f_s(H,T;L_\parallel,L_\perp) \approx -\frac{1}{V}\ln 2\cosh\left[\left(\frac{HV}{k_BT}\right)^2 + \left(\frac{L_\parallel}{L_\perp}e^{-\tilde{\sigma}_0}\right)^2\right]^{1/2} \tag{B7}$$

where the $T = 0$ surface tension enters, correctly, through the identification

$$\tilde{\sigma}_0 \equiv 2KL_\perp^{d-1}/a^{d-1} = \Sigma(T=0)A/k_BT \tag{B8}$$

in which the last part of this equation follows as in Appendix A.

Comparison with our results, as embodied in (3.14) or (3.19), shows that there is agreement only if one makes the correct correspondence $m_0(T=0) = 1$ and further accepts the identification

$$\xi_\parallel(T,L_\perp) \Rightarrow \tfrac{1}{2}L_\perp \exp(\Sigma A/k_BT) \tag{B9}$$

for $T \to 0$. However, the exact evaluation of $\xi_\parallel$ for the $d = 2$ Ising model (Appendix A) shows that the BNC result is too large by a factor $(L_\perp/a)$. [See (A10), which gives the result for $T \to 0$; when $L_\perp \to \infty$ at fixed $T > 0$ the error is of order $(L_\perp/a)^{1/2}$ but comparison in this limit is not really justifiable.] More generally, by (4.8) there are no grounds for accepting (B9) for other dimensionalities either!

The basic flaw in the BNC argument is that one is not justified in using the linearized flow equation (B3) when $L_\parallel/\xi_\parallel \gg 1$. This may be understood heuristically by noting that the nature of the renormalization group transformation is to smear out microscopic details. However, in the cylinder geometry the most probable configurations entail homogeneous domains of mean length $\xi_{\|}$ separated by roughly $L_{\|}/\xi_{\|}$ distinct interfaces: but an interface at low temperatures is a *micro*scopic structure, on the scale $\xi_{\infty} = O(a)$. Consequently the flow equation (B3), linearized about the bulk discontinuity fixed point, misses some of the needed information.

To demonstrate the deficiency of the BNC approach more concretely, consider $d = 2$ and suppose one renormalizes only down to $L'_{\perp} = 2a$ which describes an $L'_{\|} \times 2a$ "ladder" (or double-chain) with periodic boundary conditions. This system can again be solved exactly, and in place of (B6) one finds, for small T' and H',

$$f_{2a}(H', T'; L') \approx -\frac{1}{2L'} \ln 2\cosh\left[(L'/a)(4\bar{h}'^2 + e^{-8K'})^{1/2}\right] \quad \text{(B10)}$$

As a matter of fact this expression can be derived from (B6) merely by recognizing the scaling properties of an Ising strip for low field and temperature! On using (B3) as before but with $b = L_{\perp}/2a$ one finds, instead of (B7) the *different* result

$$f_s \approx -\frac{1}{V} \ln 2\cosh\left[\left(\frac{HV}{k_B T}\right)^2 + \left(\frac{2L_{\|}}{L_{\perp}} e^{-\tilde{\sigma}_0}\right)^2\right]^{1/2} \quad \text{(B11)}$$

The necessary identification of $\xi_{\|}$ again differs from the correct result for $d = 2$ by a power of $L_{\perp}$ but now a further factor of 2 is also present! One may rationalize the different answers by noting that, in effect, the BNC approach fails to differentiate factors of 2 which are true numerical constants from those equal to $L'_{\perp}/a$ which rescale nontrivially. Thus wrong factors of $L_{\perp}$ and wrong numerical factors are not surprising. In any event, an adequate treatment along the BNC lines would have to go beyond the simple linearized renormalization group relation (B3).

REFERENCES

1. M. E. Fisher, in *Critical Phenomena*, Proceedings of the Enrico Fermi International School of Physics, Vol. 51, M. S. Green, ed. (Academic, New York, 1971).
2. M. E. Fisher and M. N. Barber, *Phys. Rev. Lett.* **28**:1516 (1972).
3. M. N. Barber, in *Phase Transitions and Critical Phenomena*, Vol. 7, C. Domb and J. L. Lebowitz, eds. (Academic, New York). [in press].
4. M. P. Nightingale, *J. Appl. Phys.* **53**:7927 (1982).
5. M. E. Fisher and A. N. Berker, *Phys. Rev. B* **26**:2507 (1982).
6. B. Nienhuis and M. Nauenberg, *Phys. Rev. Lett.* **35**:477 (1975).
7. Y. Imry, *Phys. Rev. B* **21**:2042 (1980).
8. P. Kleban and C.-K. Hu, preprint, University of Maine, 1982.
9. H. W. J. Blöte and M. P. Nightingale, *Physica* **112A**:405 (1982).
10. V. Privman and L. S. Schulman, *J. Phys. A* **15**:L231 (1982).
11. V. Privman and L. S. Schulman, *J. Stat. Phys.* **29**:205 (1982).
12. H. W. J. Blöte, M. P. Nightingale, and B. Derrida, *J. Phys. A* **14**:L45 (1981).

13. J. L. Cardy and P. Nightingale, *Phys. Rev. B* **27**:4256 (1983).
14. L. Onsager, *Phys. Rev.* **65**:117 (1944).
15. M. E. Fisher, *J. Phys. Soc. Japan Suppl.* **26**:87 (1969).
16. E. Brézin, *J. Phys. (Paris)* **43**:15 (1982).
17. M. E. Fisher, in *Essays in Physics*, Vol. 4, G. K. T. Conn and G. N. Fowler, eds. (Academic Press, London, 1972), p. 43.
18. D. Jasnow and M. E. Fisher, *Phys. Rev. B* **3**:895 (1971).
19. T. D. Schultz, D. C. Mattis, and E. W. Lieb, *Rev. Mod. Phys.* **36**:856 (1964).
20. C. J. Thompson, *Mathematical Statistical Mechanics* (MacMillan, New York, 1972), Chaps. 5, etc.
21. W. J. Camp and M. E. Fisher, *Phys. Rev. B* **6**:946 (1972).
22. W. J. Camp, *Phys. Rev. B* **6**:960 (1972).
23. P. H. Kleban and G. Akinci, preprint, University of Maine, 1983.
24. A. E. Ferdinand and M. E. Fisher, *Phys. Rev.* **185**:832 (1969).
25. A. D. Bruce, *J. Phys. C* **14**:3667 (1981).
26. M. E. Fisher and W. J. Camp, *Phys. Rev. Lett.* **26**:565 (1971).
27. C. M. Newman and L. S. Schulman, *J. Math. Phys.* **18**:23 (1977).
28. M. E. Fisher and H. Nakanishi, *J. Chem. Phys.* **75**:5857 (1981).
29. H. Nakanishi and M. E. Fisher, *J. Chem. Phys.* **78**:3279 (1983).
30. R. B. Griffiths, *Phys. Rev.* **152**:240 (1966).
31. L. S. Schulman, *J. Phys. A* **13**:237 (1980).
32. K. Binder, *Phys. Rev. Lett.* **47**:693 (1981).
33. B. Widom, *J. Chem. Phys.* **43**:3982 (1965).
34. B. Widom, in *Phase Transitions and Critical Phenomena*, Vol. 3, C. Domb and M. S. Green, eds. (Academic Press, New York, 1972), p. 79.
35. M. E. Fisher, in *Collective Properties of Physical Systems*, Proc. Nobel Symp., Vol. 24, B. Lundqvist and S. Lundqvist, eds. (Academic Press, New York, 1974), p. 16.
36. M. E. Fisher, in *Renormalization Group in Critical Phenomena and Quantum Field Theory —Proceedings of a Conference*, J. D. Gunton and M. S. Green, eds. (Temple University, Philadelphia, 1974), p. 65.
37. M. N. Barber and M. E. Fisher, *Ann. Phys. (N.Y.)* **77**:1 (1973).
38. K. G. Wilson and J. Kogut, *Phys. Rep.* **12C**:74 (1974).
39. M. E. Fisher, *Rev. Mod. Phys.* **46**:597 (1974).
40. P. Nightingale, *Proc. Kon. Ned. Akad. Wet. B* **82**:235 (1979).
41. C. J. Hamer, *J. Phys. A* **15**:L675 (1982).
42. C. N. Yang, *Phys. Rev.* **85**:809 (1952).
43. M. E. Fisher, *Physica* **25**:521 (1959).
44. J. W. Essam and D. L. Hunter, *J. Phys. C* **1**:392 (1968).
45. D. S. Gaunt and G. A. Baker Jr., *Phys. Rev. B* **1**:1184 (1970).
46. G. A. Baker Jr. and D. Kim, *J. Phys. A* **13**:L103 (1980).
47. V. Privman and M. E. Fisher, *J. Phys. A* **16**:L295 (1983).
48. C. J. Hamer and M. N. Barber, *J. Phys. A* **14**:2009 (1981).
49. P. Pfeuty, D. Jasnow, and M. E. Fisher, *Phys. Rev. B* **10**:2088 (1974).
50. P. A. Rikvold, W. Kinzel, J. D. Gunton, and K. Kaski, *Phys. Rev. B* [in press].
51. K. Binder, in *Phase Transitions and Critical Phenomena*, Vol. 7, C. Domb and J. L. Lebowitz, eds. (Academic, New York) [in press].
52. T. Niemeijer and J. M. J. van Leeuwen, in *Phase Transitions and Critical Phenomena*, Vol. 6, C. Domb and M. S. Green, eds. (Academic, New York, 1976).