

SCALING, UNIVERSALITY AND RENORMALIZATION GROUP THEORY

By

Michael E. Fisher  
Baker Laboratory, Cornell University  
Ithaca, New York 14853, U.S.A.

Lectures presented at the  
"Advanced Course on Critical Phenomena"  
held during January 1982 at  
The Merensky Institute of Physics  
University of Stellenbosch, South Africa

Lecture Notes prepared with the assistance of

Arthur G. Every  
Department of Physics, University of the Witwatersrand,  
Johannesburg, South Africa  
On leave at Department of Physics, University of Illinois  
at Urbana-Champaign  
Urbana, Illinois 61801, U.S.A.

CONTENTS

1. Introduction
2. Critical Phenomena in magnets and fluids: Universality and Exponents
  - 2.1 The gas-liquid critical point
  - 2.2 Universal behavior
  - 2.3 Binary fluids
  - 2.4 Critical exponents defined precisely
  - 2.5 Specific heats
  - 2.6 The order parameter
  - 2.7 Fluid-magnet analogy
  - 2.8 Magnetic susceptibility
  - 2.9 Critical isotherm
3. Scaling
  - 3.1 Introduction: thermodynamic functions
  - 3.2 The classical phenomenological approach or Landau theory
  - 3.3 The scaling concept
  - 3.4 Scaling of the free energy
  - 3.5 Fluctuations, correlations and scattering
  - 3.6 The correlation length
  - 3.7 Decay of correlations at and below criticality
  - 3.8 Scaling of the correlation functions
  - 3.9 Anomalous or critical dimensions: General definitions and relations
4. Microscopic Models
  - 4.1 The need for models
  - 4.2 Ising model
  - 4.3 Solution of the one-dimensional Ising model
  - 4.4 The two-dimensional Ising model
  - 4.5 Ising model in three dimensions: Series expansions
  - 4.6 The n-vector spin models
  - 4.7 Continuous spin models
5. Renormalization Group Theory
  - 5.1 Preamble
  - 5.2 A renormalization group for the one-dimensional Ising model
    - 5.2.1 Spatial rescaling and spin correlations
    - 5.2.2 Unitarity
  - 5.3 Flow equations, recursion relations, and fixed points
    - 5.3.1 Linearization about a fixed point

- 5.3.2 A second variable and scaling
- 5.4 General renormalization groups
  - 5.4.1 The space of Hamiltonians
  - 5.4.2 Renormalization group desiderata
  - 5.4.3 Quasilinear renormalization groups
- 5.5 Flows, universality and scaling
  - 5.5.1 Universality, relevance and irrelevance
  - 5.5.2 Continuous flows
  - 5.5.3 The fixed point spectrum
  - 5.5.4 Scaling of the free energy and hyperscaling
- 5.6 The construction of renormalization groups
  - 5.6.1 Kadanoff's block spin renormalization group
  - 5.6.2 Niemeijer and van Leeuwen's majority rule
  - 5.6.3 Wilson's momentum shell integration

## 6. Dimensionality Expansions

- 6.1 Transformation of the Hamiltonian
- 6.2 Computing with continuous spins
- 6.3 Implementation of momentum shell renormalization
- 6.4 The Gaussian fixed point
- 6.5 The renormalization group to order  $\epsilon$
- 6.6 The n-vector fixed point
- 6.7 Some numerics
- 6.8 Further developments in brief

## ACKNOWLEDGMENTS

Appendix A The Kac-Hubbard-Stratonovich Transformation

Appendix B Details of the  $\epsilon$ -Expansion Calculation

Appendix C Dimensionality as a Continuous Variable

Appendix D Hyperscaling and Dangerous Irrelevant Variables

## BIBLIOGRAPHY

## REFERENCES

## 1. Introduction

My aim in these lectures will be to describe some of the more interesting and important aspects of critical phenomena. I will, in particular, be discussing the ideas of scaling and critical exponents and emphasizing the idea of universality. Following this I will be dealing with the microscopic formulation of statistical mechanics and certain series expansion methods that have been extensively used in the past. These are not only applicable to critical phenomena, but are useful in other areas of physics and engineering as well. The main emphasis and focus of the lectures will, however, be on the collection of rather subtle ideas which underlie renormalization group theory and its applications to critical phenomena. I will be approaching them in roughly two stages: from the microscopics will come some introductory concepts; then I will be describing the general renormalization group ideas which are essentially topological in nature. I will be aiming, in describing these general concepts, at applications beyond critical phenomena. They have for example been used to handle the Kondo problem and to study various aspects of field theories. Finally, I will discuss some of the first practical successes of the renormalization group, based on the so-called "epsilon expansions". These expansions are generated in terms of the parameter  $\epsilon = 4-d$ , where  $d$  is the spatial dimensionality of the physical system. They were some of the first sweet fruits of the renormalization group ideas!

What is the task of theory? It is worthwhile, when embarking on theory to have some viewpoint as to what theory is. There are different opinions on this subject. Some people feel the task of theory is to be able to calculate the results of any experiment one can do: they judge a theory successful if it agrees with experiment. That is not the way I look at a theory at all. Rather, I believe the task of theory is to try and understand the universal aspects of the natural world; first of all to identify the universals; then to clarify what they are about, and to unify and inter-relate them; finally, to provide some insights into their origin and nature. Often a major step consists in finding a way of looking at things, a language for thinking about things -- which need not necessarily be a calculational scheme. This aspect of renormalization group theory, which I view as very important, is underplayed in a number of articles and books on the subject. "Shapes" are aspects I often regard as important. To make an illustrative point here, the geometrical properties of the circle have been known for a long time, and we tend to take them for granted. The ratio of the circumference to the diameter is called  $\pi$ , which is only the first of many Greek letters that will be introduced in these lectures! Its value, which today we know as 3.14159265358... was, from very early times, felt to be the same for all circles; i.e., that it was a universal property. This is true if space is Euclidean; and, to a very high degree of accuracy, the space we inhabit is, indeed, Euclidean. The value of this ratio is of course of great interest. The Bible has an unambiguous statement<sup>1</sup> that the value of

$\pi$  is 3, although the people to whom that is attributed probably knew that it was not exactly equal to 3. This "Biblical" theory is the analogue to the so-called "classical" theory of critical exponents that will be referred to frequently in these lectures. We know that the Ancient Greeks already had very good inequalities for  $\pi$ . Of course, the numerical value of  $\pi$  is now known to very many decimal places indeed, and there are numerous series expansions which converge to the exact value, such as

$$\pi = \sqrt{\frac{6}{1^2} + \frac{6}{3^2} + \frac{6}{5^2} + \dots} \quad (1.1)$$

Also, there are explicit formulae that relate  $\pi$  to the other transcendental numbers, the most famous being

$$e^{i\pi} = -1. \quad (1.2)$$

In a similar way, in the theory of critical phenomena there is a set of important numbers, the critical exponents, and they are also believed to be universal in character. In these lectures evidence will be presented to show that this is so. In addition some formulae, in the form of series expansions, have been derived for these critical exponents, but, so far, only the first few terms in the expansions are known. Also, while the expansion (1.1) for  $\pi$  is convergent, the  $\epsilon$ -expansions for the critical exponents are almost certainly not convergent in general, unless they are treated in a special way. We will also see that there are a number of formulae like (1.2) which relate the various critical exponents to one another, although perhaps with not quite the mathematical rigor and generality of (1.2).

These remarks more or less sum up the attitude I will be taking towards the subject matter of these lectures.

## 2. Critical Phenomena in magnets and fluids: Universality and Exponents

### 2.1 The gas-liquid critical point

The first critical point to be discovered was in carbon dioxide. Suppose one examines<sup>2</sup> a sealed tube containing  $\text{CO}_2$  at an overall density of about 0.5 gm/cc, and hence a pressure of about 72 atm. At a temperature of about 29<sup>o</sup> C one sees a sharp meniscus separating liquid (below) from vapor (above). One can follow the behavior of liquid and vapor densities if one has a few spheres of slightly different densities close to 0.48 gm/cc floating in the system. When the tube is heated up to about 30<sup>o</sup> C one finds a large change in the two densities since the lighter sphere floats up to the very top of the tube, i.e., up into the vapor, while the heaviest

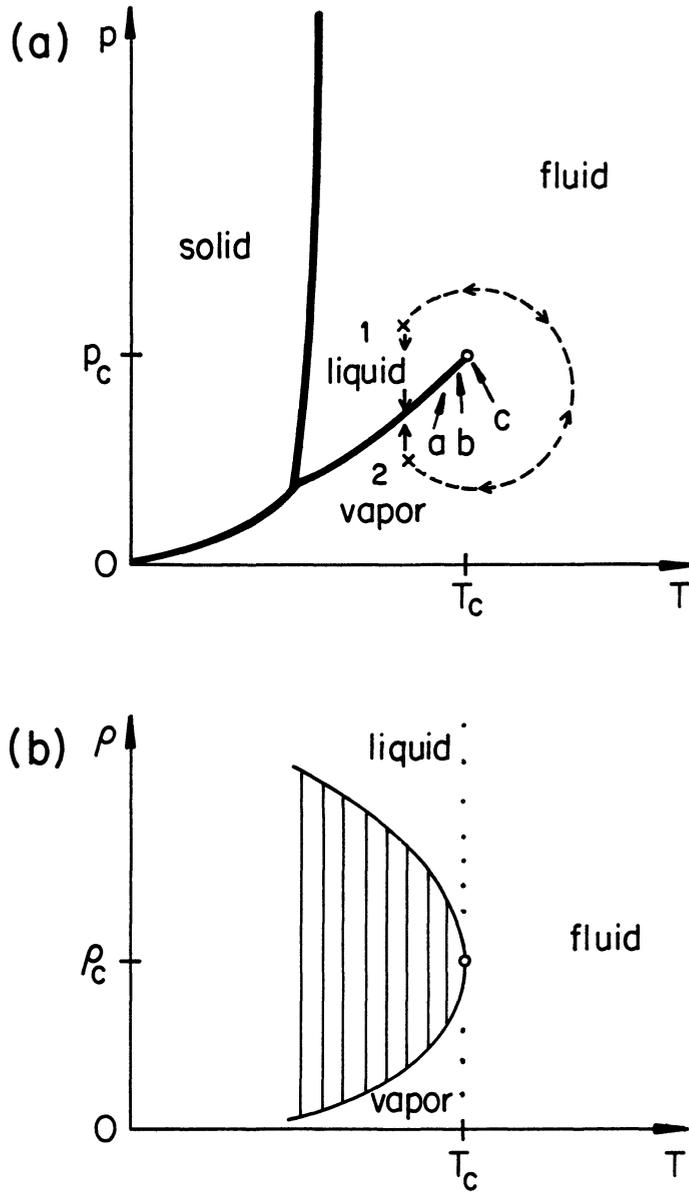


Fig. 2.1 (a)  $(p,T)$  diagram for a typical physical system;  
 (b) corresponding plot of particle number density  $\rho$  versus  $T$ .  
 The vertical "tie-lines" link coexisting liquid and vapor  
 densities, and span the region of liquid vapor coexistence.

one sinks down to the bottom of the liquid. However, a sphere of about "neutral" density (in fact "critical density") remains floating "on" the meniscus. There is, indeed, still a sharp interface between the two fluids, but they have approached one another closely in density. Further slight heating to about  $31^{\circ}\text{C}$  brings on the striking phenomenon of critical opalescence. If the carbon dioxide, which is quite transparent in the visible region of the spectrum, is illuminated from the side, one observes a strong intensity of scattered light. This has a bluish tinge when viewed normal to the direction of illumination, but has a brownish-orange streaky appearance, like a sunset on a smoggy day, when viewed from the forward direction (i.e., with the opalescent fluid illuminated from behind). Finally, when the temperature is raised a further few tenths of a degree, the opalescence disappears and the fluid becomes completely clear again. Moreover, the meniscus separating "liquid" from "vapor" has completely vanished: no trace of it remains! All differences between the two phases have gone: indeed only one, quite homogeneous, "fluid" phase remains above the critical temperature ( $T_c \approx 31.04^{\circ}\text{C}$ ).

These phenomena are best interpreted in the pressure-temperature (p,T) phase diagram shown in Fig. 2.1. The first three stages are represented by the points a, b and c on the vapor pressure curve. Note that  $T_c$  and  $p_c$  are the critical temperatures and pressures respectively at which critical opalescence is observed. As the temperature is raised further, the system follows a contour of constant overall density (the "critical isochore"). The whole process is completely reversible. Significantly, it is possible to go from liquid (point 1) to vapor (point 2) either smoothly via a route along which the properties of the fluid always change smoothly and continuously, or through the vapor pressure curve, at which a first order transition takes place with a discontinuity in density, internal energy, etc. Any point inside the shaded region of Fig. 2.1(b) corresponds to liquid and vapor coexisting with one another. As the critical point is approached the two densities,  $\rho_{\text{liq}}(T)$  and  $\rho_{\text{vap}}(T)$  become closer and closer to each other until they match at  $T = T_c$ .

## 2.2 Universal behavior

One finds that the actual variation of  $\rho_{\text{liq}}(T)$  and  $\rho_{\text{vap}}(T)$  is close to universal for gases such as argon, krypton, nitrogen, oxygen, etc., in the sense that if the temperature is normalized by the critical temperature,  $T_c$ , and the density by the critical density,  $\rho_c$ , then the data for the different gases all fit very nearly on the same coexistence curve. The shape of this coexistence curve will be one of the first objects of our investigation. The simplest curve which has the same basic shape as the coexistence curve graphed as  $T$  vs  $\rho$  is, of course, the parabola  $y = Ax^2$ . The assertion that the coexistence curve is parabolic (in the critical region) in fact represents the "Biblical" or classical theory of the

coexistence curve. At first sight, it seems to be a most natural and unprejudiced starting point. But what really is the shape of this curve near  $T_c$ ? That is the question one must ask!

To that end we introduce here a variable that will be greatly used, namely, the reduced temperature

$$t = \frac{T - T_c}{T_c}, \quad (2.1)$$

which measures the deviation of the temperature from critical in dimensionless units. Now as  $T$  approaches  $T_c$  from below, the difference between the liquid and gas densities,  $\rho_{liq}$  and  $\rho_{vap}$  respectively, is going to vanish as, we might reasonably expect, some power  $\beta$  of  $|t|$ . Thus we write

$$\rho_{liq} - \rho_{vap} \sim |t|^\beta \quad \text{as } T \rightarrow T_c^-. \quad (2.2)$$

The exponent  $\beta$  is the first of the critical exponents that will be introduced in these lectures. It is the analogue of  $\pi$  because it directly describes the shape of the coexistence curve. From the way the parabola is oriented in Fig. 2.1(b), we see that the classical or "Biblical" theory prediction is simply  $\beta=1/2$ . How does this compare with the value of  $\beta$  measured in the real world? The experiments that have been done in this connection are some of the most precise experiments ever performed in Physics. A notable example is provided by the work of Balzarini and Ohrn<sup>3</sup> who measured the coexistence curves for xenon and sulphur hexafluoride using very sensitive optical methods. These two fluids are obviously very different chemically but, nevertheless, their critical behavior is found to be essentially the same. The data on the density jump  $\Delta\rho = \rho_{liq} - \rho_{vap}$  span the range from  $t \approx 3 \times 10^{-2}$  down to  $3 \times 10^{-6}$  and on a log-log plot lie very accurately on two straight and parallel lines. This first confirms the power law behavior and then yields a value of  $\beta$  which is quite close to  $1/3$ . The precise value lies somewhere in the interval  $0.32 - 0.34$ , perhaps closer to  $0.32$ . Despite the experimental accuracy and the great range of the data one cannot, unfortunately, actually determine such critical exponents to much better than  $\pm 0.02$ . We are certain now that it is not a simple fraction, or at least not a very simple fraction like  $1/2$  or  $1/3$ ! Clearly, therefore, the "Biblical" value is quite outrageously wrong. Finally, in line with  $\pi$  being independent of the size of the circle, it is found that  $\beta$  is also quite independent of the type of fluid; the same values are found for water, a highly associated liquid, for liquid metals, and for the 'quantal liquids' helium three and four at their liquid-vapor critical points.

### 2.3 Binary fluids

Another type of system which has been much investigated is that of a mixture of two chemical compounds, say A and B, that at high enough temperatures are mutually soluble, but at lower temperatures separate out into two phases as oil separates from water, which we will call  $\alpha$  and  $\beta$  (see Fig. 2.2). There are a great many combinations that can be used: organic liquids such as aniline and cyclohexane or carbontetrachloride and perfluoroheptane are favourites because the interesting behavior occurs (under atmospheric pressure) at temperatures close to room temperature. The vapor phase is usually present, as shown in Fig. 2.2, but plays no essential role. The denser phase at the bottom could be, for example, A-rich, while the less dense one floating above it would then be B-rich. As the temperature is increased a liquid-liquid critical point or consolute point is reached and critical opalescence is exhibited just as for a one-component fluid. Beyond this point only a single, homogeneous liquid phase exists.

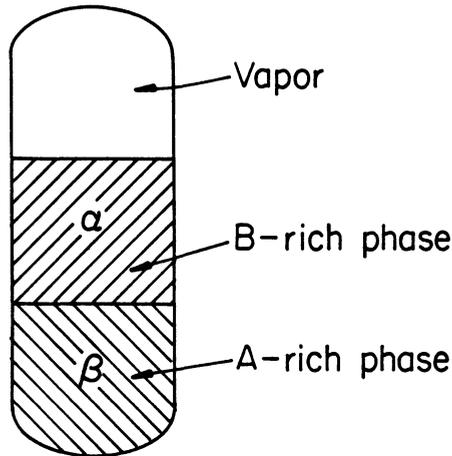


Fig. 2.2 Illustrating phase separation in a binary liquid mixture of two chemical species A and B.

Now what should one focus on instead of the density difference? We will use symbols such as  $x_A^\alpha$  to denote the mole fraction of A molecules in the A-rich phase  $\alpha$  and so on. As the critical temperature is approached from below one observes that the composition difference between  $\alpha$  and  $\beta$  phases varies as

$$x_A^\alpha - x_A^\beta \sim |t|^\beta, \quad (t \rightarrow 0^-). \quad (2.3)$$

The question is "Does  $\beta$  have the same value as before?" The answer is an

unequivocal "yes" as can be seen from experiments on very many binary fluid systems (including molten metal mixtures). An interesting comparison has been published by Sienko<sup>4</sup>. He finds, for example, that on a normalized log-log plot the coexistence curve for  $\text{CCl}_4$  and  $\text{C}_7\text{F}_{14}$  shows a form which is almost indistinguishable from that of the liquid-vapor coexistence curve for  $\text{CO}_2$ , so that  $\beta$  again lies close to  $1/3$ . Sienko and coworkers also studied the metal-ammonia systems in which alkali metals Na, Li and Ca are dissolved in  $\text{NH}_3$ . At first sight these mixtures appear to provide an exception to the  $\beta \approx 1/3$  rule. For temperatures deviating from (below)  $T_c$  by from 1% to 10% (i.e.,  $t = 0.01$  --  $0.1$ ) the coexistence curve on a log-log plot has a steeper slope than for other systems and, indeed, seems to conform to a  $\beta = 1/2$  relation as predicted by the "Biblical" theory. But accurate data that are taken closer into the critical (or consolute) point fall clearly into line with all the other systems: the slope changes quite rapidly around  $t \approx 0.007$  to  $0.009$  and decreases to yield again  $\beta \approx 1/3$ . So we are forced to accept this universality of behavior, but we learn that the universality does not extend indefinitely out of the critical region. Indeed it is really a matter of extrapolating in towards the critical point if one wants to determine the true, universal, asymptotic behavior. So when I discuss critical behavior it is always a matter of approaching close enough to the critical point. It is worthwhile to embody this point in a formal definition of a critical exponent which can then be used for more exact and rigorous theoretical arguments and analyses.

#### 2.4 Critical exponents defined precisely

Generally, when we say a function  $f(x)$  behaves like  $x^\lambda$ , or write

$$f(x) \sim x^\lambda \quad \text{as } x \rightarrow 0+, \quad (2.4)$$

it will be taken to mean that

$$\lim_{x \rightarrow 0+} \frac{\ln[f(x)]}{\ln x} = \lambda. \quad (2.5)$$

In this way we can avoid introducing a constant for the coefficient of  $x^\lambda$  as would be essential if we wrote  $f(x) \approx Ax^\lambda$  or  $f(x) \propto x^\lambda$ . At a more subtle level suppose we have a function such as

$$f(x) = A |\ln x|^\mu x^\lambda. \quad (2.6)$$

This does not vary as a simple or "pure" power law but rather has a "confluent" logarithmic singularity. From a theoretical viewpoint one can still use eqn. (2.4), and in this way one obtains a critical exponent equal to  $\lambda$ . Thus even functions of

this type with more complex singularities are covered. One of the important contributions of renormalization group theory is that it reveals the circumstances under which such logarithmic factors should be anticipated. One must always expect, of course, that over any finite range there will be some correction terms: thus even for an asymptotically pure power law one will generally have

$$f(x) = Ax^\lambda \{1 + a_0 x^\theta + \dots + a_1 x + a_2 x^2 + \dots\}, \quad (2.7)$$

where the confluent "correction" exponent,  $\theta$ , may well be less than unity (although it must be positive for the form written to make sense). On a log-log plot corrections such as these can and do actually alter the slope and lead to erroneous values for measured critical exponents. The most serious correction terms are those where  $\theta < 1$ , the smaller the value of  $\theta$  the worse the problem. In fact, values of around 1/2 are expected on theoretical grounds in many real situations. This assertion reflects another valuable contribution of the renormalization group since it has enabled us to give a sensible estimate of the exponent  $\theta$  and to explain why this sort of behavior is what one should expect in most circumstances.

There have been people in the past who have questioned whether nature really is required to conform to power law behavior near a critical point. The evidence, both experimental and theoretical, is now compelling that, apart from logarithmic factors in special cases and certain correction terms, power law behavior is the rule. One would have to be a brave scientist indeed to hold out against this conviction and this point. Nevertheless there are still those -- some would call them "cranks" -- who argue that perhaps the "Biblical" theory is still correct if one goes really close to  $T_c$ , so that  $\beta = 1/2$  after all! However, I am afraid that in science, new and more correct ideas often win out only after their opponents die or retire. Evidently many people are not as open to rational conviction by new thoughts, as might be desirable!

Another problem that arises in the handling of experimental data is that the critical temperature  $T_c$  is, of course, not known in advance. Usually one treats  $T_c$  in the expression  $t = (T - T_c)/T_c$  as a fitting parameter. When the data extends over several decades, the data close in to the transition point will sometimes be used to determine  $T_c$ , while that further out then serves to determine the critical exponents. Sometimes  $T_c$  will be determined separately from both sides in similar or distinct experiments. All in all, great care has to be exercised when interpreting even the very best data if one is not to assign misleadingly small "error" estimates to parameters such as  $T_c$ ,  $\beta$ , and the amplitudes  $A$ , etc.

## 2.5 Specific heats

In 1963 Voronel' (then in the Soviet Union) and his coworkers<sup>5</sup> made some historically important measurements of the specific heat at constant volume,  $C_V(T)$  of argon in the vicinity of its critical point. More precisely, they observed the specific heat at constant overall density along the critical isochore  $\rho = \rho_c$ . Below  $T_c$  the system will, as seen, then consist of a mixture of vapor and liquid, and the proportions of the two will actually change as the temperature is varied. So this "specific heat" actually contains a latent heat contribution. Nevertheless that is, both experimentally and theoretically, the most appropriate function to measure for the study of critical behavior. Now the "Biblical" or classical theory predicts that (I will not say "this" anymore) the specific heat merely has a jump discontinuity at the critical point, i.e.,  $C_V(T_c^-) \neq C_V(T_c^+)$ . Actually  $C_V(T_c^-) > C_V(T_c^+)$  is predicted as indicated by the dashed curve in Fig. 2.3. Voronel' was the first one to do sufficiently careful and accurate measurements to show unambiguously that this was not so! On the contrary,  $C_V(T)$  rose up smoothly but very steeply on both sides of  $T_c$  as sketched in Fig. 2.3. Asymptotically the variation has the form

$$C_V(T) \sim |t|^{-\alpha}, \quad (t \rightarrow 0^+), \quad (2.8)$$

where the specific heat exponent  $\alpha$  has a value in the region of 1/8 to 1/9 for most fluids. Because of the small value of  $\alpha$ , correction terms now assume much greater importance and make  $\alpha$  hard to determine precisely. Also one might question whether  $C_V$  does, indeed, diverge to infinity, or whether it just has a sharp spike or cusp at  $T_c$ .

On this latter question microscopic models are able to provide us with some definite guidance. These models come in various shapes and sizes: but the most famous is undoubtedly the Ising model, which I will be discussing in more detail later in these lectures. Onsager's celebrated solution of the 2-dimensional Ising model in 1944 gave the specific heat as

$$C_V(T) = A \ln|t| + \text{finite "background" terms.} \quad (2.9)$$

The singular behavior is carried mainly by the leading logarithmic term (although terms like  $t \ln|t|$  appear in the "background"). As is readily confirmed by application of the formal definition (2.5), a logarithmic divergence corresponds to the limiting case of  $\alpha \rightarrow 0^+$ . [Consider the function  $\ell_\alpha(t) = (|t|^{-\alpha} - 1)/\alpha$ .] To draw attention to the fact that the logarithm is present, this case is usually reported as

$$\alpha = 0 \text{ (log).}$$

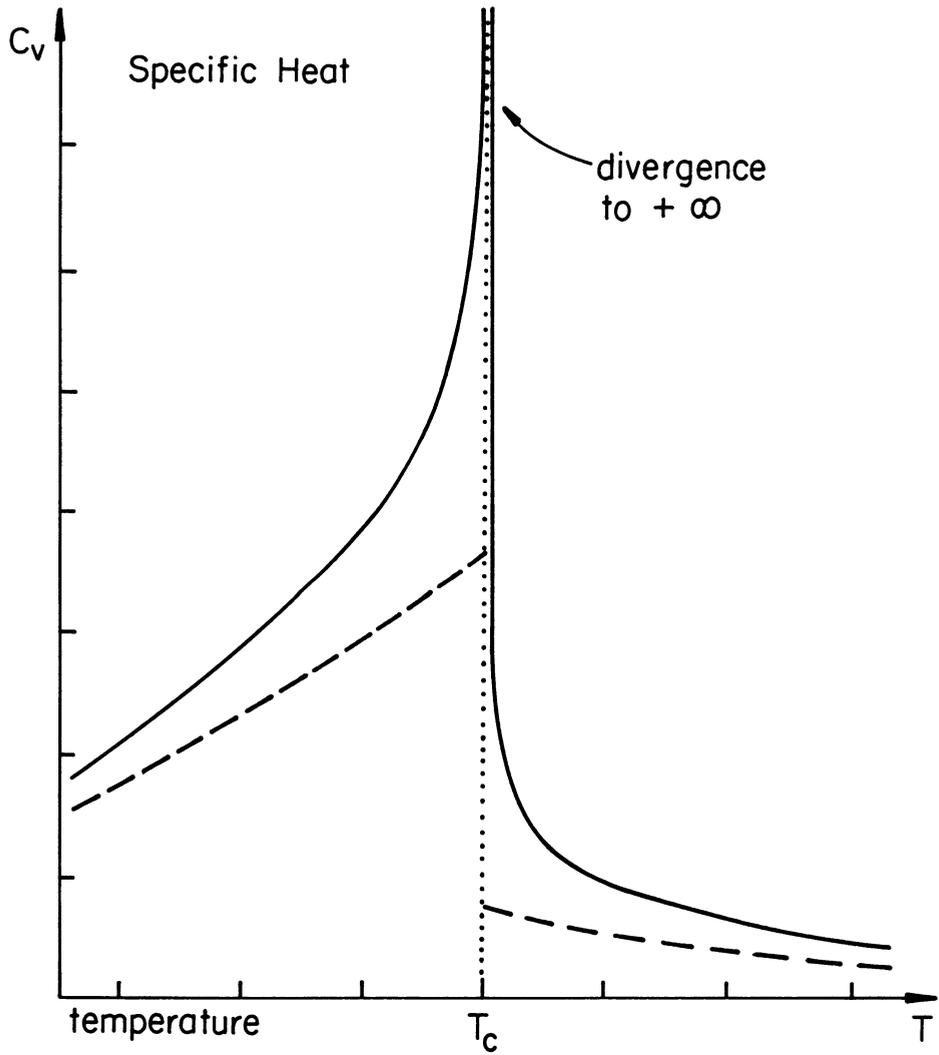


Fig. 2.3 Sketch showing the variation of the specific heat,  $C_v(T)$ , of argon and other fluids through the gas-liquid critical point. The dashed curve represents the prediction of the classical (or "Biblical") theories.

One small detail that Fig. 2.3 suggests one should take into account is that the specific heat does not mirror itself around the critical point. Thus one should, properly, define two exponents:  $\alpha'$  for  $T < T_c$  and  $\alpha$  for  $T > T_c$ . The convention is that primed exponents refer to  $T < T_c$  and unprimed to  $T > T_c$  (except where, like  $\beta$ , the definition makes sense only for  $T < T_c$ ). Nowadays it is rather well established on both experimental and theoretical grounds that  $\alpha = \alpha'$ , so the distinction is often dropped unless one has reason for being circumspect.

Modern experiments on critical specific heat obtain temperature resolutions of  $10^{-6}$  or  $10^{-7}$  in  $t$ . Some of the best experiments are those of Ahlers<sup>6</sup> on liquid helium at its lambda point,  $T_\lambda \equiv T_c \equiv 2.18$  K, where the normal fluid becomes superfluid. The transition is seen to remain sharp down to a tenth of a microdegree. More recently Lipa<sup>7</sup> has pushed the resolution still further down to only tens of nanodegrees. The specific heat seems to continue rising down to these very small deviations from  $T_c$ .

It is worthwhile asking the question at this point if, with continuing experimental refinements, one can expect to observe the specific heat continuing to diverge indefinitely close to  $T_c$ . Naturally, precautions must be taken to allow for gravity and other small disturbing factors. However, ultimately the basic theoretical answer is "No, the specific heat cannot increase without bound". The reason is that in the laboratory one would always be dealing with a finite system, with a finite number of atoms confined in a bounded region of space. A perfectly sharp phase transition can take place only in a truly infinite system, i.e., in the thermodynamic limit where the system is infinitely large in extent but its density, pressure, and all other intensive quantities are fixed and finite. However large a system is in practice, it will still be finite and, ultimately then one will reach the point where the specific heat singularity is seen to be rounded off. Experiments deliberately done on small samples certainly show these rounding effects. So in talking about a phase transition one really should always have in mind the thermodynamic limit.

The specific heat anomaly at the lambda transition in He<sup>4</sup> is now believed to be very close to logarithmic. Thus Ahlers quotes  $\alpha = \alpha' \approx -0.02 \pm 2$  (the uncertainty being in the last decimal place) signifying that  $\alpha$  is probably very slightly negative. This suggests that the specific heat does not quite diverge to  $\infty$  but rather comes up to form a sharp cusp at which point  $C(T_c)$  is finite but the slope  $(dC/dT)_c$  is infinite.

Similar behavior is also observed at magnetic phase transitions: a notable case being the specific heat of the ferromagnet nickel near its Curie or critical point,  $T_c$ . Magnetic systems are in many ways much simpler to think about theoretically because magnetic field  $H=0$  is a point of symmetry. One finds that the zero-field specific heat of nickel displays a sharp cusp, but it is much less strong than in the case of superfluid helium or some of the other fluid systems. In this

case, and that of other magnetically isotropic magnetic systems, one finds that  $\alpha$  is definitely negative although still quite small say  $\alpha = -0.10$  to  $-0.15$ , high precision being again difficult to attain.

## 2.6 The order parameter

In the case of simple fluids the parameter of apparently central interest is the density,  $\rho$ . Following Landau's general conception of phase transitions, we name this special quantity the order parameter and denote it generally as  $\Psi$ . So for single-component fluids we write  $\Psi = \rho$ . For fluid mixtures we saw that what mattered was the difference between the mole fractions,  $\Delta x$ , which measures differences in composition: so here we have  $\Psi = \Delta x$ . For superfluid  $\text{He}^4$ , the crucial theoretical concept, which embodies our understanding of superfluidity, is an effective macroscopic wave function,  $\psi = \psi' + i \psi''$ . As a wavefunction this has both real and imaginary parts. While  $\rho$  and  $\Delta x$  are both simple scalar quantities, a complex number is best thought of as a two-component vector. Thus the superfluid order parameter,  $\Psi = \psi$ , is a two-component vector which has the symmetry of a circle, i.e., can point in any direction in the complex plane. It is the phase of  $\psi$  which is in fact responsible for the existence and nature of superfluidity. In the case of ferromagnetism, there are various possibilities, but certainly it is the magnetization,  $\vec{M}$ , which should be the order parameter. In the case of a magnet like nickel, the magnetization can point freely in any direction; i.e., nickel is spatially, highly isotropic; then the magnetization can be thought of as a three-component vector  $\vec{M} = (M_x, M_y, M_z)$ .

In summary, we see that the order parameter,  $\Psi$ , has a tensorial character which may depend on the class of systems considered. Theoretically it is natural to distinguish between these various cases, and the renormalization group has enabled us to make this distinction meaningful and effective. In particular we often refer to  $n$ , the number of components of the order parameter. Then we have:

$n = 1$  for simple fluids, binary fluids, uniaxial ferromagnets, binary alloys, etc.

$n = 2$  for superfluid  $\text{He}^4$  and  $\text{He}^3 + \text{He}^4$  mixtures, XY-magnets (easy plane of magnetization).

$n = 3$  for isotropic magnets, etc.

As regards values of the critical exponents, none of which conform to "Biblical" or classical theory, there is found to be a subtle dependence on  $n$ . Specifically, one has  $\alpha(n=1) \approx 0.11$ ,  $\alpha(n=2) \approx 0.0$  and  $\alpha(n=3) \approx 0.14 \pm 4$ . Similar

slight differences are found for the critical exponent  $\beta$ , viz.  $\beta(n=1) \approx 0.32$ ,  $\beta(n=2) \approx 0.34$  and  $\beta(n=3) \approx 0.35 - 0.37$ . The  $n = 2$  value applies to XY-magnets but the corresponding superfluid order parameter is essentially inaccessible to experiment. Clearly then, the symmetry or tensorial character of the order parameter is important. The three cases described above are often referred to as Ising-like ( $n=1$ ), XY-like ( $n=2$ ), and Heisenberg-like ( $n=3$ ). Larger values of  $n$  are not just of theoretical interest; they are also required for describing real physical systems, in particular various magnetic crystals of more complex structure and symmetry.

## 2.7 Fluid-magnet analogy

The close analogies that exist between fluids and ferromagnets are worth emphasizing, even though ferromagnets have an intrinsic symmetry that makes them easier to think about. Conjugate to the order parameter,  $\Psi$ , in any thermodynamic system, is a "thermodynamic field" variable,  $h$ . In the case of fluid the pressure,  $p$ , has traditionally been treated as this conjugate variable, but often it is better to regard  $h$  as the chemical potential,  $\mu$ . The pressure  $p$  or chemical potential  $\mu$  is the variable that directly allows one to alter the density (at constant temperature). The analogous variable for a magnet should therefore be the magnetic field  $H$ , which is the variable primarily coupled to the magnetization. Fig. 2.4 illustrates clearly how far the analogy can be taken. In the case of the magnet, in the  $(h, T)$  plane, there is a line of first order transitions separating the "up" and "down" ferromagnetized states; this line ends at the critical (or Curie) point. The first order transition line is analogous to the vapor pressure curve, but differs from it in one minor respect in that it is entirely confined to the  $H=0$  or  $T$ -axis. This, of course, is a consequence of symmetry under  $H \rightarrow -H$ . In the  $(\Psi, T)$  plane there is a coexistence curve in both cases. Inside this curve the magnet breaks up into domains; this is analogous to gas-liquid coexistence in fluids. For the magnet the coexistence or "spontaneous magnetization" curve is symmetric about the  $T$ -axis while for the fluid this symmetry is apparently absent. Below  $T_c$  the order parameter variation for the fluid is given by  $\rho_{liq} - \rho_{vap} \sim |t|^\beta$  while for the magnet it is  $M_0(T) \sim |t|^\beta$ , where the spontaneous magnetization should be defined as

$$M_0(T) = \lim_{H \rightarrow 0^+} M(H, T). \quad (2.10)$$

This careful definition of  $M_0(T)$  is necessary because  $M$  takes different limiting values depending on whether  $H=0$  is approached from positive or negative values. The specific heat exponent is also defined in an analogous way for the two systems, and so on. Thus while most emphasis will be placed on magnetic systems, analogous effects and similar results hold for other types of systems in nearly all cases.

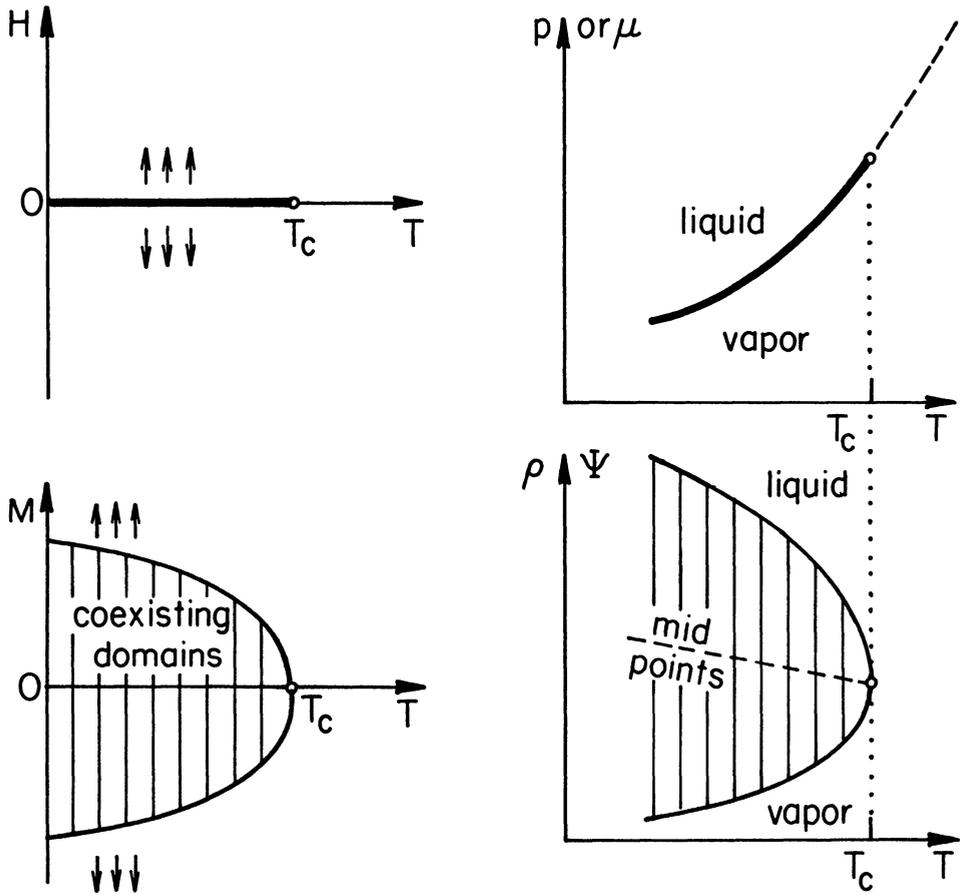


Fig. 2.4 Phase and coexistence diagrams illustrating the magnet-fluid analogy. Note magnetization corresponds to density and magnetic field to pressure or, better, chemical potential.

The question of how the perfect symmetry of the spontaneous magnetization curve is reflected in the less than fully symmetric nature of the fluid coexistence curve is a fairly subtle one. For the magnet the natural field variable to take, because of the symmetry, is  $H$ . One suspects that for a fluid the most suitable variable by analogy should be

$$h = p - p_{\sigma}'(T - T_c), \quad (2.11)$$

where  $p_{\sigma}'$  is the limiting slope of the vapor pressure line at  $T_c$ . In this way  $h$  would measure the deviation from the limiting tangent (shown dashed in Fig. 2.4), which one expects might be the analogue of the  $H=0$  symmetry axis of the magnet. This is sometimes called a scaling axis. A remarkable feature of the coexistence curve is that the line of mid points between the liquid and vapor phases is surprisingly straight. Furthermore, one can clearly define two different exponents,  $\beta_-$  and  $\beta_+$ , with respect to deviations below and deviations above critical density,  $\rho_c$ , i.e., for the vapor and liquid sides of the coexistence curve. There is no obvious (or known) symmetry between liquid and gas that should tell us a priori that these two exponents should be the same; yet to an exceedingly high degree of accuracy they are identical in value! Somehow the system builds itself an asymptotic symmetry from a Hamiltonian which does not, in the first place, possess this symmetry at all. Again, the renormalization group is able to explain how a system is able to build up a symmetry on approach to a critical point, and to decide when a symmetry can be built (or, on the contrary, when a weakly broken near symmetry of the Hamiltonian is amplified).

## 2.8 Magnetic susceptibility

Above  $T_c$  the spontaneous magnetization of a ferromagnetic material is identically zero, but magnetization can be induced by applying a magnetic field,  $H$ . Fig. 2.5 illustrates the type of isotherms observed.

The isothermal susceptibility is defined quite generally as a function of  $H$  and  $T$  by

$$\chi_T(T, H) = \left( \frac{\partial M}{\partial H} \right)_T. \quad (2.12)$$

One usually measures, and is most interested in, the so-called initial susceptibility

$$\chi_T^0(T) = \lim_{H \rightarrow 0^+} \chi_T(T, H), \quad (2.13)$$

which measures the slope of the magnetization isotherm at zero field (as shown by

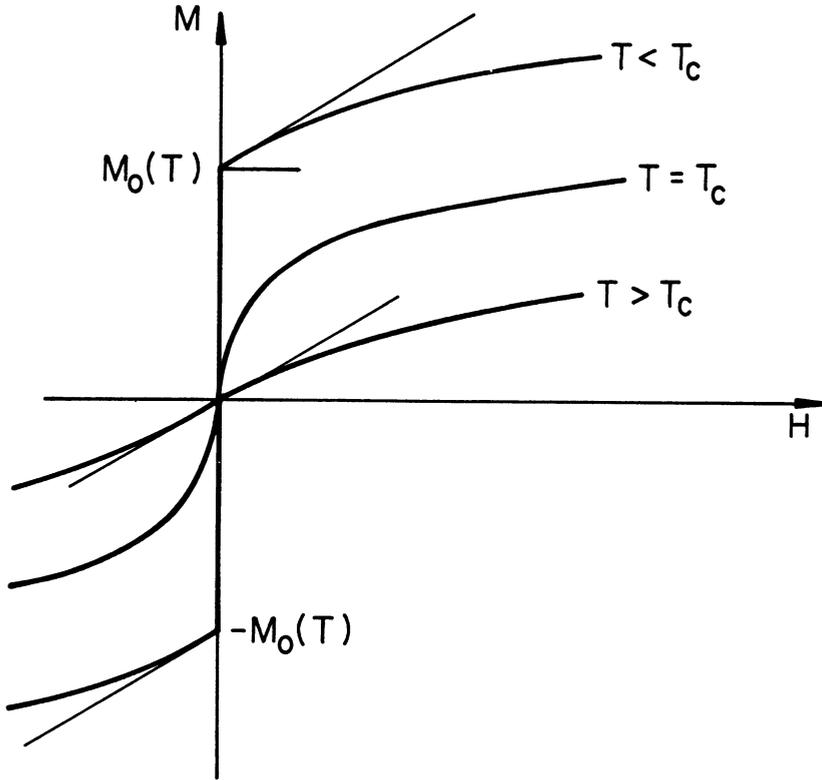


Fig. 2.5 Typical ferromagnetic magnetization curves (isotherms) above  $T_c$ , at  $T_c$  and below  $T_c$  (for a scalar, Ising-like or  $n=1$  system).

the tangents in Fig. 2.5). In practice one often drops both the superscript  $o$  and the adjective "initial" and just refers to "the susceptibility". Clearly  $\chi_T$  measures the ease of magnetizing a ferromagnet and hence is expected to grow large and, indeed, diverge at the Curie point where, after all, a ferromagnet essentially magnetizes itself! This divergence can be seen in Fig. 2.5: the slope of the critical,  $T=T_c$  isotherm is actually infinite at zero field. For theoretical purposes it is usually convenient to define the reduced susceptibility  $\chi = \chi_T/\chi_T^{\text{ideal}}$ , where  $\chi_T^{\text{ideal}}$  is the isothermal susceptibility of an ideal paramagnet (with no spin-spin interactions). Evidently,  $\chi$ , which is dimensionless, measures the enhancement in magnetic responsiveness caused by the interactions, which are, of course, responsible for the ferromagnetic critical behavior. The analogous reduced susceptibility for a fluid is  $\chi = K_T/K_T^{\text{ideal}}$ , where

$$K_T = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial p} \right)_T, \quad (2.14)$$

is the isothermal compressibility of the fluid and  $\kappa_T^{\text{ideal}} = 1/p$  is the corresponding quantity for an ideal gas. Since, as explained,  $\chi$  measures the ease with which the order parameter is changed in response to the conjugate field, it is often known as the response function. (See also the lectures by A. L. Fetter).

The divergence of  $\chi(T)$  at criticality is very strong and is characterised by an exponent  $\gamma$  defined as expected via

$$\chi \sim 1/t^\gamma, \quad (t \rightarrow 0^+; h = 0). \quad (2.15)$$

Measured values of  $\gamma$  are typically  $\gamma(n=1) \approx 1.23 \text{ -- } 1.24$ ,  $\gamma(n=2) \approx 1.31 \text{ -- } 1.32$  and  $\gamma(n=3) \approx 1.35 \text{ -- } 1.38$ . In the case of superfluid  $\text{He}^4$ , one does not know how to measure  $\chi$ : thus only the exponent  $\alpha$  can be measured (of the thermodynamic properties we have defined). As can be seen,  $\gamma$  has a small  $n$ -dependence, but in all cases deviates markedly from the "Biblical" value which is simply  $\gamma=1$ .

Below  $T_c$  the situation is more complex. Even at  $H=0$  there is a nonzero spontaneous magnetization,  $M_0(T)$ . Nevertheless, (as mentioned), one can still define the initial susceptibility as the limiting slope of the magnetization curve when  $H \rightarrow 0^+$ . The temperature dependence of  $\chi$ , so defined, provides one with the further exponent  $\gamma'$ . These last remarks apply, however, only to the Ising-like case of  $n=1$ . If  $n=2$  or  $3$ , so that a continuous (rotational) symmetry is present it can be shown theoretically, although experimentally it is not so easy to observe, that this limiting slope is infinite, so that  $\chi_T(T,H)$  diverges as  $H \rightarrow 0^+$  for  $T < T_c$  and the exponent  $\gamma'$  cannot be defined in the usual way.

## 2.9 Critical isotherm

The order parameter variation on the critical isotherm is generated by fixing the temperature precisely at  $T_c$ , varying the order field,  $h$ , and observing the change in  $\Psi$ , i.e.,  $M$  or  $\rho$  as the case may be. For a magnet one finds that for small  $H$  this variation is given by (see Fig. 2.5)

$$M(T = T_c) \sim H^{1/\delta}, \quad (H > 0, T=T_c), \quad (2.16)$$

which defines the critical exponent  $\delta$ . Values of  $\delta$  are typically:  $\delta(n=1) \approx 4.8$ ,  $\delta(n=2) \approx 4.7$ , and  $\delta(n=3) \approx 4.6$ . These should, perhaps, be regarded as more theoretical than experimental, since  $\delta$  is extremely difficult to measure accurately owing to the steepness of the critical isotherm. The classical value is  $\delta = 3$  which corresponds to a cubic curve for the critical isotherm. Of course, this is just the simplest analytic function which has the correct shape.

Naturally the critical isotherm near a fluid critical point displays completely

analogous behavior. The relation is often written in reverse form as

$$|p-p_c| \sim |\Delta\rho|^\delta, \quad (T=T_c), \quad (2.17)$$

where  $\Delta\rho = \rho - \rho_c$ , but this clearly corresponds precisely to the expected magnet-fluid analogy. Likewise, "Biblical" theory (in this case the original prophet is van der Waals) predicts  $\delta=3$ , a cubic relation, but experiment yields  $\delta \approx 4.2$  to 4.8.

### 3. Scaling

#### 3.1 Introduction: thermodynamic functions

The "Biblical" or classical theories break down completely in the region of a critical point. What then, can replace them? It turns out that the simplest phenomenological theories that come anywhere close to explaining critical behavior embody the concept of scaling. In order to make the discussion reasonably comprehensive one needs to couch it in terms of the full thermodynamics. Let us consider a ferromagnet since its symmetry allows us to make certain convenient (but inessential) simplifications. The Helmholtz free energy,  $F(T,H)$ , is associated with the basic differential thermodynamic relation

$$dF = -SdT - MdH, \quad (3.1)$$

where  $S$  is the total entropy. From this one can, by means of a Legendre transformation, generate the alternative free energy function,  $A(T,M) = F+MH$ , and it is then a simple matter to show that the basic differential relation becomes

$$dA = -SdT + HdM. \quad (3.2)$$

The magnetic field and susceptibility are obtained from  $A$  by differentiation according to

$$H = \left(\frac{\partial A}{\partial M}\right)_T \quad \text{and} \quad \chi^{-1} = \left(\frac{\partial^2 A}{\partial M^2}\right)_T. \quad (3.3)$$

Note that the susceptibility will diverge when  $T \rightarrow T_c$ , but it is intrinsically non-negative: Indeed a negative static compressibility or magnetic susceptibility is thermodynamically inconceivable. This is equivalent to the statement that the free energy  $A$  as a function of  $M$  must be a convex function: although the graph of  $A$  versus  $M$  can have a flat portion, its curvature must, otherwise, be strictly positive (See Fig. 3.1 below).

### 3.2 The classical phenomenological approach or Landau theory

The simplest type of phenomenological theory in this context derives from mean field theory; it was developed to a fine art by Landau and now frequently goes under his name. It consists, first of all, in identifying the order parameter,  $\Psi$ , (physically if this is possible but otherwise just as an abstract quantity), and then expanding the appropriate free energy as a Taylor series in powers of the order parameter. For a magnet the issue is straightforward: we have  $\Psi = M$  and the power series expansion reads

$$A(T,M) = A_0(T) + A_2(T)M^2 + A_4(T)M^4 + \dots \quad (3.4)$$

By symmetry under  $M \leftrightarrow -M$  no odd powers of  $M$  can be present. At high temperatures this expansion can be justified for all reasonable models on fully rigorous grounds, but near  $T_c$  it turns out to be dangerous! By differentiating twice one obtains the inverse susceptibility, which in zero field above  $T_c$  is thus given by

$$\chi^{-1} = 2A_2(T) \text{ for } T > T_c, \quad (H,M = 0). \quad (3.5)$$

The next assumption is that the coefficients  $A_j(T)$  can also be expanded in powers of  $t \propto (T-T_c)$  so that, in particular, we may write

$$\chi^{-1} = 2A_{2,0} + 2A_{2,1}t + O(t^2). \quad (3.6)$$

When  $T \rightarrow T_c+$  the susceptibility, by definition of  $T_c$ , diverges to infinity, so that  $\chi^{-1} \rightarrow 0$  as  $t \rightarrow 0+$ , and hence  $A_{2,0} = 0$ . The predicted behavior of  $\chi$  near  $T_c$  is thus

$$\chi \approx C/t \quad \text{as} \quad t \rightarrow 0+, \quad (H,M = 0). \quad (3.7)$$

This, of course, corresponds to  $\gamma = 1$ . The fact that this theory gives an incorrect value for  $\gamma$  can be traced directly to the unjustifiable assumption that  $A(T,M)$  can be expanded in a power series near and, indeed, at a critical point. Nevertheless this seems to be a very natural assumption of the sort which is frequently made in physics and engineering. Furthermore, it can also be shown to be the essentially inevitable outcome of any of the wide variety of more microscopically-based mean field theories that have been proposed in this and many other related contexts.

In spite of the evident shortcomings of the classical phenomenological theory, let us continue to explore its consequences by considering the effect of the term of fourth order in  $M$ . Its coefficient is

$$A_4(T) = A_{4,0} + O(t) = \frac{1}{4}u + O(t), \quad (3.8)$$

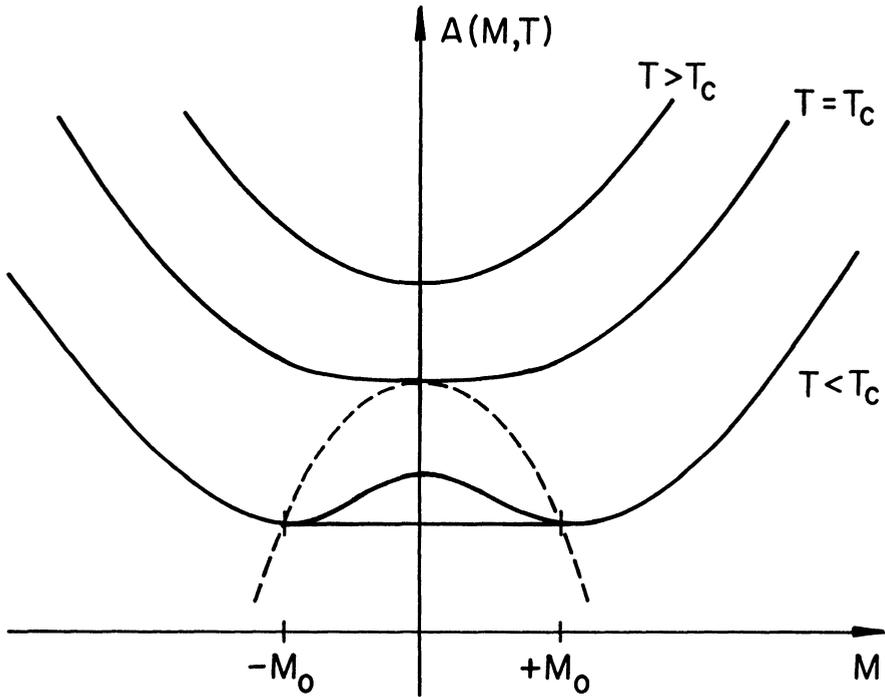


Fig. 3.1 Variation of the free energy  $A(T, M)$  according to classical phenomenological theory. The non-convex section of the isotherms for  $T < T_c$  must be "corrected" by drawing in the flat, tangential segment, so forming the "convex cover" of the underlying, approximate function.

where the replacement of  $A_{4,0}$  by  $\frac{1}{4}u$  is purely a matter of convenience. We will assume  $u > 0$  to ensure thermodynamic stability (although, in fact, the case  $u < 0$  is required for dealing with tricritical points). Let us now examine the equation of state, which is the relationship connecting  $T$ ,  $H$  and  $M$ , near  $T_c$ . It is obtained by differentiating  $A$  with respect to  $M$  and is easily seen to be of the form

$$H \approx M(ct + uM^2), \quad (3.9)$$

where we have put  $2A_{2,1} = c$  so that, from (3.5) and (3.6),  $A_2(T) \approx \frac{1}{2}ct$ . For a fluid the corresponding equation would, for example, follow from van der Waal's equation with  $M$  replaced by  $\rho - \rho_c$  and  $H$  by  $p - p_c$ .

On setting  $t = 0$ , we obtain the critical isotherm as  $H \sim M^3$  and, thence, the erroneous prediction  $\delta = 3$ . For  $T < T_c$  and  $H \rightarrow 0^-$  one obtains an equation with three roots, namely,

$$M = 0$$

and 
$$M = \pm M_0(T) \approx B|t|^{\frac{1}{2}}$$

with 
$$B = (c/u)^{\frac{1}{2}}. \quad (3.10)$$

The first root turns out to have a higher free energy than the other two (see Fig. 3.1) and therefore is of no real physical interest. The other two roots provide two equivalent states of equilibrium spontaneous magnetization. We see clearly that the predicted value of the exponent  $\beta$  is  $\frac{1}{2}$ , the incorrect classical result.

One of the difficulties of the classical theory is associated with the necessary convexity of the free energy. If one follows through in graphical terms the arguments just presented, one obtains for the variation of  $A$  as a function of  $M$  for various values of  $T$  the results shown in Fig. 3.1. Above  $T_c$  the variation predicted by (3.4) is quadratic in  $M$  for small  $M$  and obviously convex. At  $T_c$  the coefficient of the quadratic term vanishes and  $A$  has a pure fourth power dependence on  $M$ . The graph is extremely flat but still convex as it should be. Below  $T_c$ , however, the coefficient of  $M^2$  is negative, so the curve starts off at  $M = 0$  like an inverted parabola, although it is ultimately turned around by the positive quartic term. The resulting concave portion of the curve for small  $M$  is clearly unphysical, and this should be taken as an indication that the theory has gone wrong! This defect in the theory can, however, be repaired in a more-or-less ad hoc way by means of the "Maxwell construction", which essentially consists of drawing a straight line between the two minima at  $-M_0(T)$  and  $+M_0(T)$ . This process generates the so-called "convex cover" of the original  $A(M)$  plot. But what can be done about the totally incorrect values of the critical exponents that come from this theory? Can anything

be salvaged? The temptation is to somehow or other graft on the correct values! This desire brings one naturally to the idea of scaling.

### 3.3 The scaling concept

There are several ways in which the desired modifications of Landau theory can be introduced. One of the earliest and most direct approaches was that of Widom. The gist of the argument goes as follows: Consider the exponent  $\beta$ . The incorrect classical value of  $1/2$  arises from the presence of the  $M^2$  term in (3.9). Let us therefore try to patch up the theory by replacing  $M^2$  by  $M^{1/\beta}$  where  $\beta$  is now a free parameter that can be fitted to experiment. If this were all, the equation of state would thus become

$$H \approx M(ct + uM^{1/\beta}), \quad (3.11)$$

and so the spontaneous magnetization below  $T_c$  would come out correctly! Likewise, however, one might try to get the susceptibility exponent,  $\gamma$ , right by replacing  $t$  in (3.9) by  $t^\gamma$ . For  $T > T_c$  it follows that  $H \approx cMt^\gamma$  and so  $\chi \sim t^{-\gamma}$  as desired. If this modification is to apply also for negative values of  $t$  then  $t$  should obviously be replaced by  $|t|$ . This, however, is easily seen to lead one into trouble since it introduces non-analytic behavior into the equation of state everywhere on the critical isotherm  $t = 0$  (even for  $H$  or  $M$  nonzero). This has quite unphysical consequences since, in fact, the equation of state is, both theoretically and experimentally, completely free of singularities on crossing the critical isotherm away from  $H = M = 0$ . Similar problems arise in (3.11) for small values of  $M$  above  $T_c$ , where the expansion (3.4) should be valid but is not unless  $1/\beta$  is an even integer!

To avoid these problems let us rewrite (3.9) by dividing through by  $c|t|^{3/2}$  to obtain the equivalent form

$$D \frac{H}{|t|^\Delta} = \left( \frac{M}{B|t|^\beta} \right) \left\{ \pm 1 + \left( \frac{M}{B|t|^\beta} \right)^{1/\beta} \right\}, \quad (3.12)$$

where we have replaced  $3/2$  by  $\Delta$  while  $B$  and  $D$  are simply related to the original constants  $c$  and  $u$ . However, from this point on we may release  $\Delta$  from its constrained value and treat it as a second free exponent, which, hopefully, can be adjusted to get, say,  $\gamma$  correct. Now the spontaneous magnetization varies as  $|t|^\beta$  so the quantity  $M/B|t|^\beta$  can be viewed as the magnetization scaled by the spontaneous magnetization,  $M_0(T)$ . Similarly, on the left hand side of (3.12) we have the magnetic field,  $H$ , scaled by a characteristic power of the temperature, namely,  $|t|^\Delta$ . Next we notice that the full equation of state is a relation connecting  $M, T$  and  $H$  which we could express as  $M = \mathcal{M}(T, H)$ . Widom's original suggestion was that,

perhaps, when  $T \rightarrow T_c$ , and  $M$  and  $H$  become small the equation of state in general simplifies if  $M$  is replaced by the suitably scaled magnetization, namely  $M/B|t|^\beta$  and  $H$  is replaced by a suitably scaled field, namely  $DH/|t|^\Delta$ . This evidently applies to the special case (3.9) which embodies classical theory but perhaps it also holds asymptotically for the true equation of state in the critical region! More explicitly, the nature of the proposed simplification is that, in the critical region, the equation of state reduces from a function  $\mathcal{M}(T,H)$  of two variables to a function of only one variable, but which relates the two scaled variables together. In other words, we make the scaling postulate

$$\frac{M}{|t|^\beta} \approx BW\left(D\frac{H}{|t|^\Delta}\right), \quad (3.13)$$

where  $W$  is some sufficiently general function of a single argument. This assertion, the scaling ansatz, must, at this stage, be regarded purely as a guess, albeit, as we shall see, a remarkably successful guess!

In the classical theory we have  $\beta = 1/2$  and  $\Delta = 3/2$  and these values are universal for all systems: they simply arise from the integral exponents in the assumed Taylor series expansion. Additionally in classical theory, as one sees from (3.12), the full scaling function,  $W(y)$ , is also universal. Thus we may expect more generally that  $\beta$  and  $\Delta$  are universal exponents and  $W(y)$  is a universal function, even though the values will differ from their classical counterparts. On the other hand, the parameters  $B$  and  $D$ , like  $T_c$  itself, must reflect the details of the particular ferromagnet: thus they are referred to as non-universal amplitudes. The exponent  $\Delta$  is often termed the gap exponent.

Let us now examine some of the implications of this simple but, in fact, far-reaching assumption. The susceptibility for  $t > 0$  and  $H \rightarrow 0$  is given by

$$\chi \propto \left(\frac{\partial M}{\partial H}\right)_{T,H=0} \approx |t|^{\beta-\Delta} BDW'(0), \quad (3.14)$$

where  $W'(0)$  must just be some number. Since, by definition, we have  $\chi \sim t^{-\gamma}$ , we see that

$$\Delta = \beta + \gamma. \quad (3.15)$$

This shows how  $\Delta$  should be chosen to give the right value of  $\gamma$ . Otherwise it tells us nothing new.

To find a new result let us look at the critical isotherm,  $T = T_c$ , for which purpose the limit  $t \rightarrow 0$  must be studied. In this limit the scaled magnetic field evidently diverges since

$$y = D\frac{H}{|t|^\Delta} \rightarrow \infty \quad \text{as} \quad t \rightarrow 0. \quad (3.16)$$

In the spirit of the enterprise let us then assume that  $W(y)$  also varies as some power when  $y$  becomes large, i.e., suppose

$$W(y) \approx W_\infty y^\lambda \quad \text{as } y \rightarrow \infty, \quad (3.17)$$

where  $W_\infty$  and  $\lambda$  are constants. It follows that

$$M \approx |t|^\beta D^\lambda B W_\infty H^\lambda / t^{\lambda \Delta}. \quad (3.18)$$

When  $t \rightarrow 0$ , the temperature variable should drop out of this expression since  $M$  then becomes a function of  $H$  only; consequently we must demand  $\lambda \Delta = \beta$  which fixes the exponent  $\lambda$  as

$$\lambda = \beta / \Delta. \quad (3.19)$$

Thus there is, in reality, no free choice of  $\lambda$ ! Moreover, from (3.17) we now see that  $M \sim H^\lambda$ ; but, by definition we have  $M \sim H^{1/\delta}$  for  $T = T_c$ . Thus we conclude  $\delta = 1/\lambda$  and hence

$$\delta = \frac{\Delta}{\beta} = 1 + \frac{\gamma}{\beta}. \quad (3.20)$$

This novel equation relating the three exponents  $\beta$ ,  $\gamma$ , and  $\delta$  is known as Widom's relation. It is our first nontrivial scaling law or, simply, exponent relation.

In a similar way, by integrating  $M = -(\partial F / \partial H)_T$  to obtain the free energy  $F(T, H)$  and then differentiating with respect to  $T$  one derives expressions for the entropy and specific heat and hence establishes the so-called Essam-Fisher relation

$$\alpha' + 2\beta + \gamma' = 2. \quad (3.21)$$

A little further investigation using the fact that there must be no singularities as one crosses the critical isotherm,  $t = 0$ , at nonzero  $H$  or  $M$  reveals that one must also have

$$\alpha = \alpha', \quad (3.22)$$

for the specific heats above and below  $T_c$  and

$$\gamma = \gamma', \quad (3.23)$$

for the susceptibilities. Evidently there are four relations connecting the six exponents  $\alpha$ ,  $\alpha'$ ,  $\beta$ ,  $\gamma$ ,  $\gamma'$  and  $\delta$ , and so only two of them can be independent: this is a striking prediction, by now verified many times experimentally!

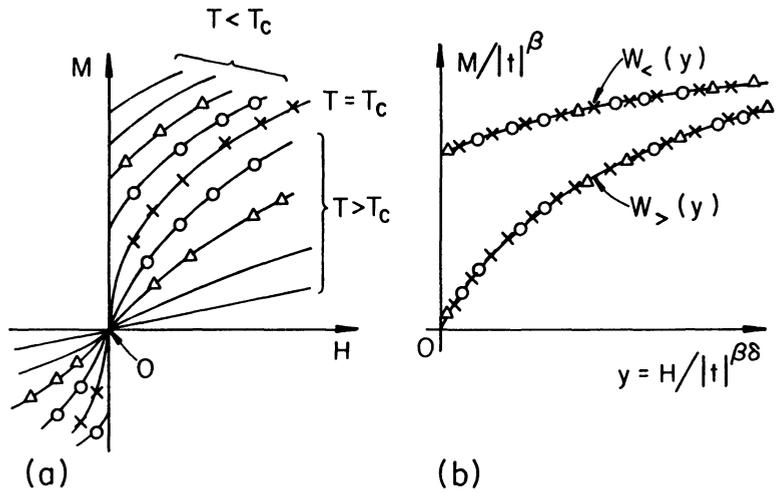


Fig. 3.2 (a) A schematic plot illustrating equation of state data, i.e.,  $M$  versus  $H$  isotherms, for a ferromagnet through the critical region; (b) a scaled plot of the same data illustrating the "collapse" of the data onto a single scaling function  $W(y)$ , with two branches  $W_>(y)$  and  $W_<(y)$  corresponding to  $t \gtrless 0$ .

The success of scaling can be illustrated graphically by replotting equation of state data for magnets, fluids, etc. in scaled form. Thus consider Fig. 3.2(a) where  $M$  versus  $H$  isotherms are sketched for a ferromagnet. This data may be replotted in scaled form as  $M/|t|^\beta$  versus  $y = H/|t|^{\beta\delta}$  (recalling that  $\Delta = \beta\delta$  by the scaling laws) for appropriate choice of the exponents  $\beta$  and  $\delta$ , which might be determined separately from the spontaneous magnetization curve and critical isotherm. Scaling is confirmed if one observes, as in fact is found,<sup>8</sup> a "collapse" of the data for the different isotherms onto a common locus, which represents the scaling function  $W(y)$ . Actually in this representation one finds two branches, asymptotically matching as  $y \rightarrow \infty$ , corresponding to  $W_>(y)$  and  $W_<(y)$ , the scaling function for  $T \gtrless T_c$ . When the procedure is repeated for different magnets one finds similar results with, indeed, the same scaling function<sup>†</sup> up to different scaling amplitudes  $B$  and  $D$ . The scaling function that emerges for fluids is, likewise, the same for all fluids, and furthermore it agrees, as do the exponents, with that found for magnets!<sup>†</sup>

To sum up then, the scaling postulate proves to be a remarkably successful guess. Our theoretical task from here on is to set scaling theory in a broader context, to explain why it works, and to ask if we can actually calculate the exponents and, also, the scaling functions. The renormalization group approach provides many of the explicit answers and, further, explains the circumstances under which scaling can break down and how it fails.

### 3.4 Scaling of the free energy

It is useful at this stage to recapitulate by taking a somewhat different approach to scaling, and to be a little more precise. Specifically we will again use the symbol " $\sim$ " to mean "behaves like" and take the symbol " $\approx$ " to mean "asymptotically equal to" i.e., if  $f(x) \approx g(x)$  as  $x \rightarrow 0$ , the ratio,  $f(x)/g(x)$  approaches unity when  $x \rightarrow 0$ . As before, the discussion will be couched in magnetic language but, as previously emphasized, the same types of behavior are to be found in many other systems if one merely identifies the analogous quantities properly.

In the critical region the free energy,  $F(T,H)$ , will have a singular part which embodies the leading critical behavior. Let  $\Delta F$  be the deviation of the free energy from its value at the critical point with other non-singular contributions (the "background" terms) also subtracted off. We define a normalized or reduced free energy by

---

<sup>†</sup>We restrict attention here to uniaxial,  $n = 1$  or Ising-like magnets. For isotropic,  $n = 3$ , Heisenberg-like or XY,  $n = 2$  magnets of different symmetry the exponents differ slightly, (see later below) and, necessarily, the scaling functions must also differ slightly for these distinct "universality classes."

$$f_{\text{singular}} = \frac{-\Delta F}{k_B T V} \quad (3.24)$$

Division by  $k_B T$  produces a dimensionless quantity which, however, is still proportional to the size of the system; thus, we have also divided by the volume,  $V$ , in order to obtain an intensive quantity which contains the bulk thermodynamics. The dimensions of  $f$  are thus inverse volume or number density. As we will be working in the thermodynamic limit,  $f$  is independent of  $V$  (which we suppose becomes infinite through a sequence of domains of reasonable shape).

In the previous section scaling was introduced via the equation of state in (3.13). If we integrate  $M(T,H)$  with respect to  $H$ , this leads to the free energy which (after background subtraction) will be similarly scaled. Alternatively, we could introduce a scaling postulate directly for the free energy. In this way, the scaling ansatz becomes the assertion

$$f_{\text{sing.}}(T,H) \approx A_0 |t|^{2-\alpha} Y\left(D \frac{H}{|t|^\Delta}\right) \quad \text{as } t, H \rightarrow 0, \quad (3.25)$$

where  $A_0$  and  $D$  are non-universal scaling amplitudes which depend on the details of the system. The first,  $A_0$ , sets the scale of the free energy while  $D$  sets the scale of the magnetic field. As before, there appear two universal exponents  $\alpha$  and  $\Delta$ . A technical point that arises here concerning the scaling function  $Y(y)$  was already alluded to before: specifically, the universal function  $Y(y)$  should really be considered in two parts:  $Y_>(y)$  for  $t > 0$  and  $Y_<(y)$  for  $t < 0$ . These two parts must match analytically as  $y \rightarrow \infty$ , but to pursue that point here would be unnecessarily distracting. The reason for writing the power of the temperature prefactor in (3.25) as  $2 - \alpha$  is to get the specific heat exponent correct, as is easily seen. Let us set  $H = 0$  and normalize the scaling function by setting  $Y(0) = 1$ , which we may do because of the presence of the factor  $A_0$ . Recalling that entropy is given by  $S = -(\partial F/\partial T)_{H=0}$  it follows that the singular part of the entropy varies as

$$\Delta S(T) \propto \frac{\partial f}{\partial t} \approx A_1 t^{1-\alpha}, \quad (H=0). \quad (3.26)$$

The internal energy behaves similarly. The specific heat then follows as

$$C(T) = T \left(\frac{\partial S}{\partial T}\right) \propto \frac{\partial^2 f}{\partial t^2} \approx A_2 t^{-\alpha}. \quad (3.27)$$

In these expressions  $A_1$  and  $A_2$  are amplitudes proportional to  $A_0$ . (Note that the variation of the prefactor  $T$  in the definition of  $C(T)$  is smooth and so does not affect the critical behavior of the specific heat. For  $t > 0$  the symbol  $C$  referred to here could be subscripted either M or H since in zero field ( $H=0$ ) one has  $C_M = C_H$ : this is a special feature resulting from the symmetry of a simple

ferromagnet. More generally, one should consider  $C_M(T)$  or, for a fluid system,  $C_V(T)$  and so on.

The reader should sketch the variation of the zero field entropy  $S(T)$ , noting that  $S$  is monotonic, and also continuous through the critical point,  $T = T_c$ , but exhibits a vertical tangent there, varying in the vicinity as  $\pm |t|^{1-\alpha}$  where, as is typically true, we have supposed  $\alpha$  has a positive, albeit small value. Moreover, the internal energy and many other quantities 'driven' by the critical behavior, such as the resistance, exhibit precisely the same form near  $T_c$ . Thus it is not these quantities themselves but rather their temperature derivatives which diverge at  $T_c$  (or, if  $-1 < \alpha < 0$  exhibit a sharp cusp there).

The equation of state  $M = \mathcal{M}(T, H)$  is obtained by differentiating with respect to  $H$ . This yields

$$M = -\left(\frac{\partial F}{\partial H}\right)_T \approx \frac{\partial f}{\partial H} \approx A_0 D |t|^{2-\alpha-\Delta} Y'\left(D \frac{H}{|t|^\Delta}\right). \quad (3.28)$$

The reason for calling  $\Delta$  the gap exponent can now be seen. Each successive differentiation with respect to  $H$ , to form  $\chi = (\partial M / \partial H)$ ,  $\chi_2 = (\partial^2 M / \partial H^2)$ , etc. changes the exponent of the  $|t|$  prefactor by the constant decrement  $\Delta$ . For  $T < T_c$  and  $H \rightarrow 0$  the scaling function  $Y'$  will approach a nonzero constant value and so, as before,  $M_0(T) \approx B |t|^{2-\alpha-\Delta}$ . But since, by definition,  $M_0(T) \sim |t|^\beta$ , it follows that

$$\beta = 2 - \alpha - \Delta. \quad (3.29)$$

Adopting  $M/|t|^\beta$  as the scaled magnetization, we see from (3.28) that this is a function only of the scaled magnetic field,  $y \propto H/|t|^\Delta$ , thus recapturing the original scaling postulate (3.13).

To obtain the critical isotherm we let  $t \rightarrow 0$ , and in line with the arguments used in the previous section, assume that  $Y(y) \approx y^{\lambda+1}$  when  $y \rightarrow \infty$ . The choice  $\lambda = \beta/\Delta$  ensures that  $|t|$  cancels out when  $t \rightarrow 0$ . In this way, repeating the details for the sake of completeness, we obtain

$$M \sim t^\beta \left(\frac{H}{t^\Delta}\right)^{\beta/\Delta} \sim H^{\beta/\Delta}. \quad (3.30)$$

But since  $M \sim H^{1/\delta}$  we conclude

$$\Delta = \beta\delta, \quad (3.31)$$

as before [see (3.20)].

Finally, the susceptibility is given by  $\chi_T = (\partial M / \partial H)_T$  and for the reduced susceptibility above  $T_c$  we find

$$\chi \propto \frac{\partial^2 f}{\partial H^2} \approx A_0 D^2 t^{2-\alpha-2\Delta} \approx Ct^{-\gamma}, \quad (3.32)$$

so that we obtain the scaling relation

$$\gamma = -2 + \alpha + 2\Delta. \quad (3.33)$$

By combining (3.29), (3.31) and (3.33) one readily establishes the various scaling relations

$$\alpha + \beta(1 + \delta) = 2, \quad (3.34)$$

$$\alpha + 2\beta + \gamma = 2, \quad (3.35)$$

and 
$$\Delta = \beta + \gamma = \beta\delta. \quad (3.36)$$

Quite clearly, the classical values of the exponents, viz.  $\alpha = 0$ ,  $\beta = 1/2$ ,  $\gamma = 1$  and  $\delta = 3$  satisfy these relations! Even before the full advent of scaling, Rushbrooke had shown on rigorous thermodynamic grounds that, because of the convexity of the free energy, the exponent inequality

$$\alpha' + 2\beta + \gamma' > 2, \quad (3.37)$$

was a thermodynamic necessity. Note that this is a rigorous result that does not depend on any assumption as scaling theory does. Similarly, Griffiths later proved the inequality

$$\alpha' + \beta(1 + \delta) > 2, \quad (3.38)$$

corresponding to (3.34). Evidently, then, scaling theory certainly does not conflict with thermodynamics even though it asserts that the rigorous inequalities hold as equalities. Nor, however, can the scaling laws be obtained by pure thermodynamic arguments although quite a few theorists have been tempted to think so and to try to demonstrate it! Occasional reports in the past of measured values of critical exponents violating the above inequalities have all proved to be poorly founded (which is just as well, since otherwise a violation of the Second Law of Thermodynamics would have been observed!)

It seems that, at least as far as systems belonging to the same symmetry class are concerned, the critical exponents are universal quantities satisfying the scaling laws. Similarly, the scaling function  $Y$  is a universal function only of the scaled field  $y$  for such systems. However, as mentioned before, one does expect some change in  $Y$  as  $n$ , the number of components of the order parameter and  $d$ , the spatial

dimensionality, are varied. So we can write  $Y = Y(y;n,d)$ . To emphasize this point, note, as will be shown, that there are good grounds for believing that classical theory is correct when  $d > 4$ . To see how the scaling function depends on  $d$ , consider the behavior of the (zero-field) susceptibility above and below criticality. We can write

$$\begin{aligned} \chi &\approx C^+/t^Y && \text{as } T \rightarrow T_c^+, \\ &\approx C^-/|t|^Y && \text{as } T \rightarrow T_c^-, \end{aligned} \quad (3.39)$$

where the amplitude ratio,  $C^+/C^-$ , should be universal but, clearly, depends on the particular form of the scaling function. Within Landau theory it is an easy exercise to prove  $C^+/C^- = 2$  (which is, indeed, universal). We can accept this for  $d > 4$ , but for the Ising model ( $n = 1$ ) and  $d < 4$ , we find  $C^+/C^- \approx 5.03$  for  $d = 3$  while for  $d = 2$  one knows the exact universal value  $C^+/C^- = 37.693562\dots$

### 3.5 Fluctuations, correlations and scattering

What is the 'cause' of the failure of mean field theory and Landau theory? Why do they yield wrong exponents and wrong scaling functions? The short answer is "Because they neglect fluctuations". To understand the significance of this piece of now conventional wisdom and to explore further striking critical phenomena that provide a key to the renormalization group approach, let us study fluctuations in the critical region and introduce the correlation and scattering functions which serve to quantify them and to describe relevant observations.

Much can be learned about criticality by scattering radiation --- light, x-rays, neutrons, etc. --- off the system of interest. In a standard scattering experiment, a well-collimated beam of light, or other radiation, with known wavelength,  $\lambda$ , is directed at the sample, fluid, magnetic crystal, etc., and one measures the intensity,  $I(\theta)$ , of the light scattered at an angle  $\theta$  away from the "forward" direction of the main beam. The radiation undergoes a shift in wave vector,  $\underline{k}$ , which is simply related to  $\theta$  and  $\lambda$  by

$$|\underline{k}| = \frac{4\pi}{\lambda} \sin \frac{1}{2} \theta. \quad (3.40)$$

The scattered intensity  $I(\theta)$  is determined by the fluctuations in the medium. If the medium were perfectly uniform (i.e., spatially homogenous) there would be no scattering at all! If one has in mind light scattering from a fluid, then the relevant fluctuations correspond to regions of different refractive index and, hence, of particle density  $\rho(\underline{R})$ . For neutron scattering from a magnet, fluctuations

in the spin or magnetization density are the relevant quantities, and so on. We need to study the normalized scattering intensity  $I(\theta, T, H, \dots)/I^{\text{ideal}}(\theta)$ , where  $I(\theta; T, H, \dots)$  is the actual scattering intensity observed at an angle  $\theta$ , which will normally depend on such factors as temperature, magnetic field, etc. while  $I^{\text{ideal}}(\theta)$  is the scattering that would take place if the individual particles (spins, etc.) doing the scattering could somehow be taken far apart so that they no longer interacted and thus were quite uncorrelated with one another. Now this normalized scattering intensity is proportional to the fundamental quantity

$$\hat{G}(\underline{k}) = \int d\underline{R} e^{i\underline{k} \cdot \underline{R}} \tilde{G}(\underline{R}), \quad (3.41)$$

which represents the Fourier transform of the appropriate real space correlation function  $G(\underline{R})$  (of density-density, spin-spin, etc.)

As the critical point of a fluid or fluid mixture is approached one observes enormously enhanced values of the scattering, especially at low angles, corresponding via (3.40) and (3.41), to long wavelength density fluctuations in the fluid. In the immediate critical region the scattering is so large as to be visible to the eye, particularly through the phenomenon of critical opalescence. This behavior is not, however, limited to fluids. Thus if, for example, one scatters neutrons from iron in the vicinity of the Curie point one likewise sees a dramatic growth in the low-angle neutron scattering intensity as sketched in Fig. 3.3. (With neutrons care must be taken to ensure that the total elastic scattering is observed since the proportionality of  $I(\theta)$  to  $\hat{G}(\underline{k})$  holds only if inelastic scattering processes can be neglected.) As can be seen, for small angle scattering there is a pronounced peak in  $I(\theta, T)$  as a function of temperature, and this peak approaches closer and closer to  $T_c$  as the angle is decreased. Of course, one could never actually observe zero-angle scattering directly, since this would mean picking up the oncoming main beam, but one can extrapolate to zero angle. When this is done one finds, in fact, that the zero-angle scattering  $I(0, T)$ , actually diverges at  $T_c$ . This is the most dramatic manifestation of the phenomenon of critical opalescence and is quite general, being observed whenever the appropriate scattering experiments can actually be performed.

In order to understand these effects we need to examine the correlation function for the relevant quantity, which, in general, is the locally fluctuating order parameter,  $\Psi(\underline{R})$ , for the transition in question. Thus  $\Psi(\underline{R})$  could, for instance, describe how the spin varies from lattice site to lattice site in a magnetic crystal. The overall spatial average of this quantity is what was previously referred to as the (total) order parameter,  $\Psi$ . We will define the correlation function  $G_{\Psi\Psi}(\underline{R})$ , or, for brevity, just  $G(\underline{R})$  by

$$G_{\Psi\Psi}(\underline{R}) = \langle \Psi(\underline{0})\Psi(\underline{R}) \rangle - \langle \Psi(\underline{0}) \rangle \langle \Psi(\underline{R}) \rangle. \quad (3.42)$$

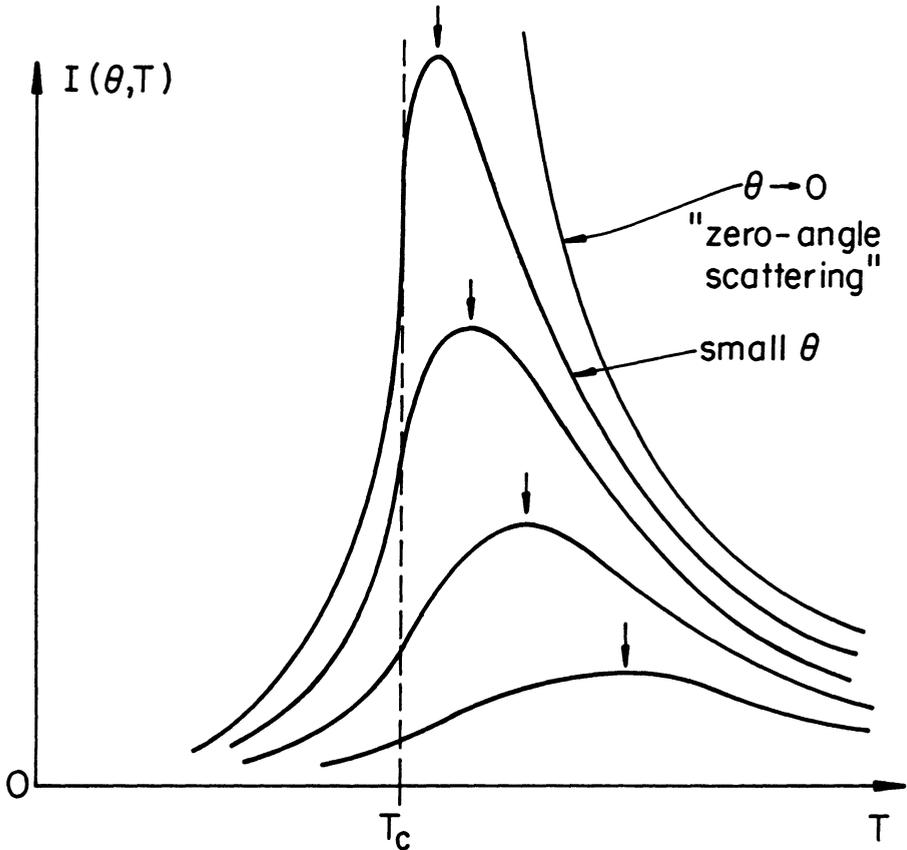


Fig. 3.3 Schematic plot of the elastic scattering intensity of neutrons scattered at fixed angle,  $\theta$ , from a ferromagnet, such as iron, in the vicinity of the Curie or critical point. The small arrows mark the smoothly rounded maxima (at fixed  $\theta$ ) which actually occur above  $T_c$  (in contrast to classical and most mean field theories which yield a nonanalytic maximum at  $T = T_c$ ).

We will always presuppose a macroscopically large system, so that there is translational symmetry. Likewise, we suppose that inhomogeneous effects due to gravity, etc. can be ignored. This means that the two average quantities,  $\langle \Psi(\underline{Q}) \rangle$  and  $\langle \Psi(\underline{R}) \rangle$  will be equal to one another and to the overall, bulk, thermodynamic order parameter  $\Psi$ . We may thus let

$$\delta\Psi(\underline{R}) = \Psi(\underline{R}) - \langle \Psi \rangle, \quad (3.43)$$

represent the deviation or fluctuation of  $\Psi$  about its uniform mean value; then it is a matter of simple algebra to show that the correlation function directly measures the fluctuations since

$$G(\underline{R}) = \langle \delta\Psi(\underline{Q}) \delta\Psi(\underline{R}) \rangle. \quad (3.44)$$

For simplicity we will often assume an isotropic system so that  $G$  is a function of  $R$  rather than  $\underline{R}$ .

### 3.6 The correlation length

If one thinks of a lattice of spins above  $T_c$  in zero field,  $H = 0$ , one has  $\langle \Psi \rangle = 0$  by symmetry. A ferromagnetic exchange coupling between neighboring spins then tends to align the spins parallel to one another whereas thermal energy works to randomize them. Thus at high temperatures one expects the spin-spin correlation function,  $G(\underline{R})$ , to fall off rather rapidly with the distance,  $R$ , separating the spins, whereas at lower temperatures the spins should become correlated with each other over longer and longer distances, the correlation function then decaying more slowly with  $R$ .

What should the law of correlation decay be? On fairly general grounds one can show\* that away from  $T_c$  the correlation function should fall off exponentially with  $R$  for large distances, i.e., that the leading behavior is given by

$$G(R) \sim e^{-R/\xi} \quad \text{as } R \rightarrow \infty, \quad (3.45)$$

where  $\xi$  is a quantity that has the dimensions of length, and is thus called the correlation length. It evidently tells us the scale on which the correlations decay. At high temperatures  $\xi$  will be just a few angstroms, but near a critical point it becomes very large. This ties in well with our earlier comments on critical opalescence, since if  $\xi$  becomes comparable with the wavelength of the

---

\*One must assume that the interactions themselves are of finite range or decay rapidly.

radiation, the medium will then contain fluctuations or inhomogeneities on that scale, and this will give rise to strong low angle scattering i.e., to critical opalescence.

There is another, very general theoretical route that tells us that  $\xi$  must become large near a critical point. This utilizes the fluctuation-susceptibility relation which reads

$$\hat{G}(0;T,H) = \int d\mathbf{R} G(\mathbf{R};T,H) = \chi(T,H). \quad (3.46)$$

For simplicity we consider here only the magnetic case. Note that  $\hat{G}(0;T,H)$  is the limiting value as  $k \rightarrow 0$  of the Fourier transform of  $G(\mathbf{R};T,H)$ . It thus depends only on  $T$  and  $H$  and is therefore a thermodynamic function: via statistical mechanics one finds it is just the reduced susceptibility,  $\chi(T,H) = k_B T \chi_T / m^2$ , where  $m$  is the magnetic moment per spin. Now when  $T \rightarrow T_c$  for  $H = 0$ , we know that  $\chi$  diverges; somehow this divergence must also come out of the integral in (3.46). Since  $G(\mathbf{R})$  is a bounded function it cannot, itself, diverge [In the case of  $S = 1/2$  spins one has  $G(\mathbf{R}) < \frac{3}{4}$ ]; thus a divergence of the integral can only mean that  $G(\mathbf{R})$  decays very slowly when  $T = T_c$ , certainly more slowly than an exponential. Consequently we are forced to conclude that  $\xi(T)$  diverges to infinity when  $T \rightarrow T_c$ . The variation of  $\xi$  near  $T_c$  can, naturally, be described by

$$\xi(T) \sim 1/t^\nu, \quad (M = 0), \quad (3.47)$$

where for three-dimensional systems the new exponent,  $\nu$ , has values around 2/3. This contrasts with the classical prediction  $\nu = 1/2$  (which follows from an extension of phenomenological, Landau theory to inhomogeneous situations). More concretely one has  $\nu \approx 0.63$  for Ising-like ( $n = 1$ ) systems, particularly fluids, increasing to  $\nu \approx 0.70$  for Heisenberg-like systems. For the two-dimensional Ising model the divergence of  $\xi(T)$  was established by Onsager along with his original calculation of the zero-field free energy which revealed the logarithmic divergence of the specific heat; his results yield  $\nu = 1$ . In experiments on fluids such as carbon dioxide, the correlation length has been measured down to  $t \approx 10^{-4}$  or  $10^{-5}$  by when  $\xi$  is thousands of angstroms in magnitude. The divergence of the correlation length is one of the crucial clues to our general understanding of critical phenomena; the renormalization group approach, in particular, focuses on the behavior of the correlation length.

### 3.7 Decay of correlations at and below criticality

At  $T_c$  the correlation length is infinite. If it were not, then the integral in (3.46) would necessarily converge and be finite: then  $\chi$  would be bounded at  $T_c$

which is certainly not the case! Thus precisely at  $T = T_c$  the correlation function,  $G(R)$  cannot be an exponential function of  $R$ . Moreover, it must, in general, still decay to zero and one should, in fact, anticipate an "algebraic" or inverse power law form such as

$$G_c(R) \approx \frac{D}{R^{d-2+\eta}} \quad \text{as } R \rightarrow \infty, \quad (3.48)$$

(where the nonuniversal amplitude  $D$  should not be confused with our previous use of this symbol). The reason for writing the decay exponent in this rather special way is that  $G(R)$  frequently appears, as in (3.46), in volume integrals of the form

$$\int d\tilde{R} G(\tilde{R})X(\tilde{R}) \propto \int_0^\infty R^{d-1}G(R)X(R) dR,$$

so that the  $d$  drops out. Evidently  $\eta$  is a new critical exponent which describes how  $G(R)$  behaves at  $T_c$ . Its numerical value is always rather small and in classical theory, which, as already mentioned, will be found to apply for spatial dimensionalities  $d > 4$ , one has  $\eta = 0$ . (Of course this result provides another good reason for writing (3.47) in the form given.) For real three-dimensional systems one finds  $\eta \approx 0.03$  to  $0.06$ , but it proves to be a very difficult parameter to measure reliably and accurately in experiments. For  $d = 2$  Ising-like systems the theoretical value is  $\eta = 1/4$ ; this can even be confirmed by experiments on (effectively) two-dimensional systems. Since  $\eta$  is in all cases small, the integral in (3.46) necessarily diverges and the susceptibility is indeed infinite at  $T_c$ .

Beneath  $T_c$  there is a subtlety that has to be taken into account. The correlation function in general now exhibits long range order, i.e.,  $G(\tilde{R})$  does not decay to zero as  $R \rightarrow \infty$  but rather approaches a nonzero value, say  $G(\infty)$ . This appearance of long range order is, in fact, one of the notable characteristics of most phase transitions. In a magnet the zero field spin-spin correlation function  $\langle S(Q)S(\tilde{R}) \rangle$  when  $R \rightarrow \infty$  becomes proportional to  $[M_0(T)]^2$ , the square of the spontaneous magnetization. Via the scattering theory this leads to a magnetic Bragg peak in the scattering of strength proportional to  $[M_0(T)]^2$ . In systems such as antiferromagnets, where  $\Psi(\tilde{R})$  is a "staggered magnetization", this provides a means of measuring the spontaneous order which would, otherwise, be inaccessible to experimental observation. If one subtracts the limiting value,  $\langle S(0)S(\infty) \rangle$ , from the correlation function one obtains a net correlation function which again decays to zero. In Ising-like systems there is then also an exponent  $\nu'$  for the correlation length beneath  $T_c$ . Experimentally, one finds  $\nu' \approx \nu$  and theoretically, according to scaling, the two exponents should be exactly the same. Experimentally the scattering intensity  $I(\theta) \equiv I(\tilde{k})$  provides us with the information needed to determine  $\xi$ . It is not hard to show that  $\hat{G}(\tilde{k})$ , which we recall is essentially proportional to  $I(\tilde{k})$ , is an even function of  $k$  which, rather generally, can be expressed in the form

$$\frac{1}{\hat{G}(k)} = \frac{1}{\chi} [1 + \xi^2 k^2 + O(k^4)], \quad (T > T_c). \quad (3.49)$$

In a so-called Ornstein-Zernike analysis one thus plots  $1/\hat{G}(k)$  [or  $1/I(k)$ ] in the critical region versus  $k^2$ . The data for small  $k$  (such that  $ka \lesssim 0.1$  where  $a$  is a typical molecular dimension) usually fall close to a straight line whose intercept with the  $k^2 = 0$  axis determines the susceptibility  $\chi(T)$ . As  $T \rightarrow T_c$  this intercept falls to zero but the successive isotherms remain more or less parallel on the Ornstein-Zernike (or OZ) plot. The reduced slopes evidently serve to determine  $\xi(T)$ . Close to  $T_c$  the plots in the case of very good experiments show a slight downward curvature: this is an indication of a nonzero and positive value of the exponent  $\eta$ . Thus at  $T = T_c$  we have, by (3.48), the power law decay  $1/R^{d-2+\eta}$ , and on Fourier transformation this yields

$$\hat{G}_c(k) \approx \frac{\hat{D}}{k^{2-\eta}}, \quad (3.50)$$

asymptotically for small  $k$ . On an Ornstein-Zernike plot the curvature of the critical isotherm thus measures  $\eta$ . It should be stressed, however, that since  $\eta$  has such a small value, it is difficult to measure this curvature unambiguously: extensive data are needed and careful corrections for multiple scattering and other extraneous effects are called for. Nevertheless, a small positive value is definitely established.

### 3.8 Scaling of the correlation functions

Our treatment of correlation functions has evidently introduced two new exponents,  $\nu$  and  $\eta$ . Are these independent of each other? Are they independent of the thermodynamic exponents  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ ? Or are all the exponents somehow linked together? Let us see what the idea of scaling has to say in this context. Accordingly, with no loss of generality we write the correlation function and its Fourier transform as

$$G(\underline{R}; T, H) = \frac{\mathcal{D}(\underline{R}; T, H)}{R^{d-2+\eta}} \quad \text{and} \quad \hat{G}(\underline{k}; T, H) = \frac{\hat{\mathcal{D}}(\underline{k}; T, H)}{k^{2-\eta}} \quad (3.51)$$

which serves to pull out the critical point behavior. Now we expect (or hope!) that  $G$  and  $\hat{G}$  will exhibit some simplified behavior as  $T \rightarrow T_c$ . Scaling means that there should be some reduced description, some compression or collapse of the multivariable data. Thus the dependence of the correlation functions on three variables might, perhaps, reduce to a dependence on only two, properly scaled, variables. The behavior at  $T_c$  has been extracted in terms of the functions  $\mathcal{D}$  and  $\hat{\mathcal{D}}$ . In line with our previous application of scaling, it is thus natural to

postulate that  $\mathcal{D}$  and  $\hat{\mathcal{D}}$  are functions not of three variables, but only of two scaled variables. We saw before that the scaled magnetic field had to take the form  $H/|t|^\Delta$ ; there is no good reason to expect this to be changed in any way for the correlation functions. However, the length  $R$  should now also be replaced by a scaled length and, likewise  $k$  by a scaled wave number. We have of course, already identified a characteristic length for the problem: this is the correlation length  $\xi$ . Since  $\xi$  diverges at criticality it is reasonable to guess that it is the only length that really matters in the critical region. We conclude that the natural scaled length is  $R/\xi \approx R|t|^\nu$  and that the appropriately scaled wave number is  $k\xi \approx k|t|^{-\nu}$ . Accepting this we can, asymptotically, replace  $\mathcal{D}$  and  $\hat{\mathcal{D}}$  by scaling functions  $D$  and  $\hat{D}$  to obtain the scaling postulates

$$G(\tilde{R}; T, H) \approx \frac{D(R|t|^\nu; H/|t|^\Delta)}{R^{d-2+n}}, \quad (3.52)$$

and, quite equivalently under Fourier transformation,

$$\hat{G}(\tilde{k}; T, H) \approx \frac{\hat{D}(k/|t|^{-\nu}; H/|t|^\Delta)}{k^{2-n}}, \quad (3.53)$$

where, for simplicity, we have left out the nonuniversal scaling amplitudes needed for full normalization if we wish to explicitly exhibit the expected universality of  $D$  and  $\hat{D}$ .

Notice that the only other lengths that could conceivably play any role are the interatomic spacings or the atomic and molecular diameters on scales, say,  $a$ . But near the critical point all such lengths become extremely small compared to the range of the correlations, and so, being "overwhelmed", become unimportant to the long wavelength behavior of the fluctuations measured by  $G$  and  $\hat{G}$ . This, indeed, gives us some insight as to why there should be universality. Different fluids are found to have the same critical exponents and scaling functions. The same thing applies to magnets (if they have the same symmetry number  $n$ ). Where does this universality come from? Clearly the only important differences between different fluids can be traced to the shapes and short range interactions of their constituent molecules, i.e., to differences on a scale of a few angströms. Near the critical point, fluctuations are taking place on the scale of  $10^3 \text{ \AA}$  and beyond, so differences on a scale of a few  $\text{\AA}$  are "washed out" or "averaged over". Thus one can understand, in an intuitive way, universality as a consequence of the fact that the correlation length becomes very large so that the important "effective interactions" no longer take place on an atomic scale but rather on a semi-macroscopic scale set by  $\xi$ . On this level the microscopic differences do not matter and one obtains universality. As systems move away from criticality and the correlation length becomes smaller, the differences start to matter. These intuitive ideas, formulated most clearly in the first place mainly by Kadanoff, are capitalized upon and made more concrete in Wilson's development of renormalization group theory.

To see how the correlation function scaling hypothesis implies connections between the critical exponents, let us examine the fluctuation integral in (3.46) that leads to the zero-field susceptibility above  $T_c$ , namely,

$$\begin{aligned}\chi &= \int G(\underline{R}; T, H = 0) d^d R, \\ &\approx C_d \int_0^\infty \frac{D(Rt^\nu; 0)}{R^{d-2+\eta}} R^{d-1} dR.\end{aligned}\quad (3.54)$$

We have been explicit here in dealing with a volume integral in  $d$  dimensions; because of the assumed (asymptotic) isotropy this reduces to an integral over the radius with the factor  $C_d R^{d-1}$  representing the surface area of a  $d$ -dimensional sphere: for future reference the relevant coefficient is

$$C_d = 2\pi^{d/2} / \Gamma(\frac{1}{2}d).$$

A change of variable to the scaled combination  $x = Rt^\nu$  transforms the integral to the form

$$\begin{aligned}\chi &\approx C_d t^{-(2-\eta)\nu} \int_0^\infty D(x) x^{1-\eta} dx \\ &\approx (\text{const.}) t^{-(2-\eta)\nu}.\end{aligned}\quad (3.55)$$

Comparing with the definition  $\chi \sim t^{-\gamma}$  we discover the new scaling relation

$$\gamma = (2-\eta)\nu, \quad (3.56)$$

which relates  $\nu$  and  $\eta$ . The classical exponent values  $\gamma = 1$ ,  $\eta = 0$  and  $\nu = 1/2$  obviously satisfy this relation. Experimentally, also, this relation checks very well. If it is accepted, it actually provides the best method of measuring the elusive exponent  $\eta$ !

Our theory at this stage is what might be called "three-exponent scaling", since from only three exponents, say  $\alpha$ ,  $\Delta$  and  $\nu$ , one can obtain all the other exponents for both thermodynamic and correlation functions. Notice that all the exponent relations so far encountered have no explicit dependence on the dimensionality,  $d$  (even though the actual values of the critical exponents themselves do depend on  $d$ ). There is, however, an important further exponent relation which does involve  $d$  explicitly: this we consider now. The argument we use may, perhaps, be regarded as not very plausible but it does lead to the desired

result, and other arguments are not much more convincing! Following Kadanoff we start by noting from (3.24) and (3.25) that the exponent  $\alpha$  makes its appearance in the singular part of the free energy in the following way:

$$f_{\text{sing.}} = \frac{-\Delta F}{k_B T V} \approx t^{2-\alpha}, \quad (H = 0), \quad (3.57)$$

and that, as remarked, the dimensions of  $f$  are those of reciprocal volume or  $1/L^d$  (where  $L$  is a length). Now  $f_{\text{sing.}} \rightarrow 0$  when  $t \rightarrow 0$  and so the relevant "critical volume" is evidently diverging, i.e., there is a significant length which is diverging! But, as we have argued, there should be only one important length in the critical region, namely, the correlation length,  $\xi$ , which moreover is also diverging as  $t \rightarrow 0$ . This suggests the identification

$$f_{\text{sing.}} \sim \frac{1}{\xi^d} \sim \frac{1}{(t^{-\nu})^d}. \quad (3.58)$$

Then comparing with (3.25) yields the new relation

$$d\nu = 2 - \alpha, \quad (3.59)$$

which explicitly involves the dimensionality. This is called a hyperscaling relation to emphasize the fact that it goes beyond and cannot be derived from the ordinary scaling relations for exponents.

Notice now that this extra relation means that just two exponents, say  $\Delta$  and  $\nu$  can be used to predict all the others, i.e., we have achieved a "two-exponent scaling theory". On combining the hyperscaling relation (3.59) with various other exponent relations one can easily derive the further hyperscaling relation

$$2-\eta = d \frac{(\delta-1)}{(\delta+1)}. \quad (3.60)$$

Again, following Buckingham and Gunton, one can show by rigorous statistical mechanical arguments that this relation must (for most systems) be satisfied rigorously as an inequality, namely,

$$2-\eta < d \frac{(\delta-1)}{(\delta+1)}. \quad (3.61)$$

Once again, then, we see that scaling comes in as the borderline of a general physical inequality. Notice that the experimental observation  $\delta < 4.8$  for  $d = 3$  implies, via this inequality,  $\eta > 0.034$ .

A peculiar feature of the hyperscaling relations is that the classical exponent values do not satisfy them unless  $d = 4$ ! To check this, substitute  $\delta = 3$  and  $\eta = 0$  in (3.61) and  $\alpha = 0$  and  $\nu = 1/2$  in (3.59). This fact also serves to demonstrate

that the hyperscaling relations have a rather different status than the other scaling relations. However, they hold exactly in the two-dimensional Ising model, and renormalization group theory is able to show why the hyperscaling relations are to be expected fairly generally, why they hold for  $d < 4$  but break down for  $d > 4$ , and by what mechanisms they can be expected to fail when they do. [See Appendix D.]

As discussed, the critical exponents themselves depend not only on  $d$  but also on  $n$ , the symmetry of the order parameter; so far, however, no general exponent relations have been discovered in which  $n$  appears explicitly. On the other hand, for certain classes of problem there are special relations between exponents in dimension  $d$  for one type of system and exponents in dimensions  $d + 1$  and  $d + 2$  for different types of problem! Many of these relations also owe their genesis to renormalization group ideas.

One might mention, in closing our phenomenological discussion of scaling, that the scaling relations can also be obtained by making certain assumptions concerning the asymptotic homogeneity character of the functional relationship between thermodynamic variables. The formalism is elegant and the end results are the same, but this approach tends to obscure the physics of the situation, which is that near a critical point each important quantity has a natural scale or size. When these natural scales are used, a reduced, universal description emerges. At the critical point itself all the temperature and field scales vanish (or diverge) so that one is left with spatial self-similarity. The fluctuations of the order parameter, for example, look statistically the same on all length scales if the magnitude of  $\Psi(\underline{r})$  is rescaled appropriately. Likewise for the energy-energy fluctuations, etc., which, in the interests of simplicity, we have not discussed. These and general aspects of scaling theory are summarized briefly in the following subsection (which need not, however, be studied in order to follow the balance of these lectures).

### 3.9 Anomalous or critical dimensions: general definitions and relations

We present here, in note form, a summary<sup>10</sup> of exponent definitions and scaling relations which emphasizes the correlation functions or, more properly, the cumulants or "connected correlation functions", for a general set of critical operators (or local densities)  $A(\underline{r})$ ,  $B(\underline{r})$ , ... .

For operators or local variables  $A(\underline{r})$ ,  $B(\underline{r})$ , ...  $N(\underline{r})$ , conjugate to fields  $h_A$ ,  $h_B$ , ... including:

the <u>order parameter</u>	$\Psi(\underline{r})$	<u>conjugate to field</u> $h \equiv h_\Psi$ ,
the <u>energy density</u>	$\mathcal{E}(\underline{r})$	<u>conjugate to field</u> $t = h_{\mathcal{E}} \approx \Delta T/T_c$ ,
<u>anisotropy energy</u>	$\mathcal{Q}(\underline{r})$	<u>conjugate to field</u> $g = h_{\mathcal{Q}}$ , etc.,

the general Cumulant is

$$\mathcal{K}_{AB\dots N}(\underline{r}_1, \dots, \underline{r}_{n-1}) = \langle A(\underline{Q})B(\underline{r}_1)C(\underline{r}_2)\dots N(\underline{r}_{n-1}) \rangle_{\underline{\epsilon}} \quad (3.62)$$

with

$$\mathcal{K}_A = \langle A \rangle, \mathcal{K}_{AB}(\underline{r}) = \langle A(\underline{Q})B(\underline{r}) \rangle - \langle A(\underline{Q}) \rangle \langle B(\underline{r}) \rangle, \text{ etc.} \quad (3.63)$$

The general scaling hypothesis for critical operators A, B, ... is then

$$\mathcal{K}_{AB\dots}(\underline{r}_j; t) \approx \frac{D(\underline{r}_1/R, \dots, \underline{r}_{n-1}/R; Rt^\nu)}{\omega_A + \omega_B + \dots + \omega_N - \omega^*} \quad \text{for } r_1, \dots, r_{n-1} \gg a, \quad (3.64)$$

and  $0 < t \ll 1$ , other fields being  
at their critical values.

Here the  $\omega_A, \omega_B, \dots$  are the critical (or anomalous) dimension of A, B, ..., while  $\omega^*$  may be called the anomalous dimension of the vacuum. Note that we write t in place of |t|, which is generally needed, merely to reduce the complexity of the formulae.)

Hyperscaling means generally

$$\omega^* = 0, \quad (3.65)$$

as predicted by formal renormalization group analysis (see later); as found in the Ising model (n=1) for d=2, and in the spherical model (n=∞) for d<4; but as violated at critical points for d > 4: see Appendix D. (Note that Kadanoff, in a notation adopted by many authors, writes  $x_A, x_B$ , etc. in place of  $\omega_A, \omega_B$ , etc. but also assumes hyperscaling and so sets  $\omega^* = 0$ .)

Thermodynamic scaling i.e., the scaling of the free energy  $f \equiv f_{\text{sing.}} = -\Delta F/k_B T V$ , obtained by integration of the cumulants and is expressed by

$$f(t, h, \dots, h_A, \dots) \approx t^{2-\alpha} Y\left(\frac{h}{t^\Delta}, \dots, \frac{h_A}{t^{\phi_A}}, \dots\right) \quad (3.66)$$

where (see further below)

$$2 - \alpha = d^*/\lambda_{\underline{\epsilon}} \equiv (d - \omega^*) / (d - \omega_{\underline{\epsilon}}) \quad (3.67)$$

and

$$\Delta = \phi_\psi = \frac{\lambda_\psi}{\lambda_{\underline{\epsilon}}} \equiv \frac{d - \omega_\psi}{d - \omega_{\underline{\epsilon}}}, \quad (3.68)$$

while the "crossover exponent" is defined by

$$\phi_A = \lambda_A / \lambda_{\underline{\epsilon}}, \quad (3.69)$$

where we have introduced

$$d^* = d - \omega^*, \quad (3.70)$$

and the general scaling eigenvalues or complementary exponents,  $\lambda_A$ , through

$$\lambda_A + \omega_A = \lambda_B + \omega_B = \dots = d. \quad (3.71)$$

(Kadanoff uses  $y_A, y_B$ , etc. in place of  $\lambda_A, \lambda_B, \dots$ ). Hyperscaling implies, of course,  $d^* = d$  and other relations such as

$$d\nu = 2\beta + \gamma', \quad \beta = \frac{1}{2}(d-2+\eta)\nu, \quad (3.72)$$

etc, which can be found from the following general

Exponent relations:

Correlation exponents

$$\xi \sim t^{-\nu}, \quad \xi \sim h_A^{-\nu/\lambda_A}, \quad \nu = \nu_{\xi}; \quad (3.73)$$

(unindicated fields set to their critical values)

$$\nu_A = 1/\lambda_A \equiv 1/(d-\omega_A); \quad (3.74)$$

$$G_c(\underline{r}) = \langle \Psi(\underline{Q})\Psi(\underline{r}) \rangle_c \sim 1/r^{d-2+\eta}; \quad \eta = 2\omega_{\Psi} + 2 - d - \omega^*; \quad (3.75)$$

$$\lambda_{\xi} = d - \omega_{\xi} = 1/\nu, \quad \omega_{\Psi} = d - 1 + \frac{1}{2}\eta - \frac{1}{2}(2-\alpha)/\nu, \quad \omega^* = d - (2-\alpha)/\nu. \quad (3.76)$$

Thermodynamics

$$\chi_{AA} = \frac{\partial^2 f}{\partial h_A^2} = \int d\underline{r} \langle A(\underline{Q})A(\underline{r}) \rangle_c \sim h_A^{-\alpha_A}; \quad (3.77)$$

where

$$2 - \alpha_A = (d - \omega^*) / (d - \omega_A) = d^* / \lambda_A, \quad (\alpha = \alpha_{\xi}), \quad (3.78)$$

while for first order cumulants

$$\langle A \rangle \sim t^{\beta_A}, \quad (\beta = \beta_{\Psi}); \quad \langle A \rangle \sim h_A^{1/\delta_A}, \quad (\delta = \delta_{\Psi}), \quad (3.79)$$

with

$$\beta_A = (\omega_A - \omega^*)/\lambda_{\xi}, \quad \delta_A = \lambda_A/(\omega_A - \omega^*) = 1/(1 - \alpha_A). \quad (3.80)$$

For the cross-susceptibility one has

$$\chi_{AB} = \frac{\partial^2 f}{\partial h_A \partial h_B} = \int d\xi \langle A(Q)B(\xi) \rangle_{\mathcal{G}} \sim t^{-\gamma_{AB}}, \quad (\gamma = \gamma_{\Psi\Psi}) \quad (3.81)$$

$$\gamma_{AB} = (d - \omega_A - \omega_B + \omega^*)/(d - \omega_{\xi}). \quad (3.82)$$

Crossover (at multicritical points):

$h_A$  scales as  $t^{\phi_A}$  with

$$\Delta = \phi_{\Psi} = \beta + \gamma = \beta\delta, \quad (3.83)$$

and

$$\phi_A = \lambda_A/\lambda = \nu/\nu_A = (d - \omega_A)/(d - \omega_{\xi}). \quad (3.84)$$

General relations

$$\alpha + \beta_A + \beta_B + \gamma_{AB} = 2, \quad \phi_A = \beta_A + \gamma_{AA} = \beta_A \delta_A. \quad (3.85)$$

#### 4. Microscopic Models

##### 4.1 The need for models

The traditional approach of theoreticians, going back to the foundation of quantum mechanics, is to run to Schrödinger's equation when confronted by a problem in atomic, molecular or solid state physics! One establishes the Hamiltonian, makes some (hopefully) sensible approximations and then proceeds to attempt to solve for the energy levels, eigenstates and so on. However, for truly complicated systems in

what, these days, is much better called "condensed matter physics," this is a hopeless task; furthermore, in many ways it is not even a very sensible one! The modern attitude is, rather, that the task of the theorist is to understand what is going on and to elucidate which are the crucial features of the problem. For instance, if it is asserted that the exponent  $\alpha$  depends on the dimensionality,  $d$ , and on the symmetry number,  $n$ , but on no other factors, then the theorist's job is to explain why this is so and subject to what provisos. If one had a large enough computer to solve Schrödinger's equation and the answers came out that way, one would still have no understanding of why this was the case! Thus the need is to gain understanding, not just numerical answers: that does not necessarily mean going back to Schrödinger's equation which, in any case, should be really regarded just as an approximation to some sort of gauge field theory. So the crucial change of emphasis of the last 20 or 30 years that distinguishes the new era from the old one is that when we look at the theory of condensed matter nowadays we inevitably talk about a "model". As a matter of fact even Schrödinger's equation and gauge field theories themselves are just models of the physical world, albeit pretty good ones as far as we can presently judge!

We should be prepared to look even at rather crude models, and, in particular, to study the relations between different models. We may well try to simplify the nature of a model to the point where it represents a "mere caricature" of reality. But notice that when one looks at a good political cartoon one can recognize the various characters even though the artist has portrayed them with but a few strokes. Those well chosen strokes tell one all one really needs to know about the individual, his expression, his intentions and his character. So, accepting Frenkel's guidance,<sup>11</sup> a good theoretical model of a complex system should be like a good caricature: it should emphasize those features which are most important and should downplay the inessential details. Now the only snag with this advice is that one does not really know which are the inessential details until one has understood the phenomena under study. Consequently, one should investigate a wide range of models and not stake one's life (or one's theoretical insight) on one particular model alone. Nevertheless, one model which, historically, has been of particular importance and which has given us a great deal of confidence in the phenomenological descriptions of critical exponents and scaling presented earlier deserves special attention: this is the so-called Ising model. Even today its study continues to provide us with new insights.<sup>12</sup>

## 4.2 Ising model

This model is absolutely the simplest model of a many body system! First of all we regard space as divided up into a lattice of cells of volume  $v_0$ , each represented by a single lattice point. The easiest lattice to think about in two dimensions is the square lattice but, following our resolution, we should at least look also at some other types, such as the triangular lattice (see Fig. 4.1). At each lattice site we allow just two possible microscopic states: in the language of a ferromagnet we place an Ising "spin",  $s$ , on each site. To distinguish spins on different sites I will usually label the spins with the position vectors  $\underline{R}, \underline{R}', \dots$  of the lattice sites or with site indices  $i, j, \dots$ : thus  $s_{\underline{R}}$  or  $s_i$ , etc.

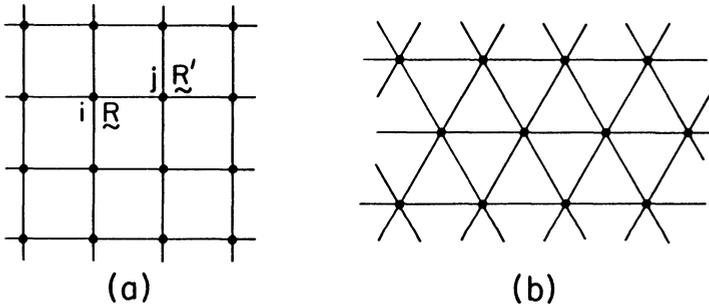


Fig. 4.1 Two dimensional lattices: (a) the plane square lattice of coordination number 4; (b) the plane triangular lattice with coordination number 6.

As implied, an Ising spin is permitted to take just two values which are expressed numerically or symbolically as

$$\begin{aligned} s_i &= +1, (\uparrow, \text{"up"}), \\ &= -1, (\downarrow, \text{"down"}). \end{aligned} \quad (4.1)$$

This two-valued variable is a similar but somewhat simpler entity than the quantum mechanical spin variable,  $\vec{S}$ , for total spin  $S = 1/2$ , whose  $z$ -component can take the two values  $S^z = \pm 1/2$ . One must notice that the Ising model also constitutes a model for a fluid, albeit the very simplest one, namely, a lattice gas. In this model we replace continuum space by the lattice of sites and suppose that the atoms or molecules can sit only on the sites. Since two atoms cannot easily be forced on top of one another, only two possibilities are contemplated at each site: either there is an atom present or there is not. Thus one can obviously establish a one-to-one correspondence between an Ising magnet and a lattice gas in which each 'down' spin,  $\downarrow$ , represents an occupied site and each 'up' spin,  $\uparrow$ , represents a vacant site: pictorially we have:

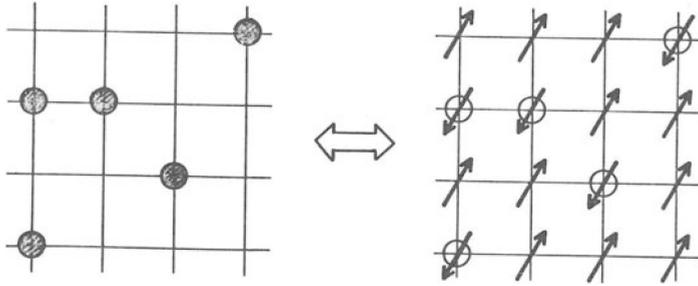


Fig. 4.2

lattice gas

Ising ferromagnet

Similarly a direct correspondence exists between an Ising ferromagnet and the simplest models of a binary alloy or of a binary fluid whose composition is the important local physical variable. One can conveniently adopt the convention that an A-atom (molecule, or ion) is represented by an 'up' spin,  $\uparrow$ , while a B-atom then corresponds to a 'down' spin,  $\downarrow$ .

Mathematically all of these situations are precisely analogous: there is only one problem to solve! An Ising spin is, clearly, just a scalar and hence the Ising model is the prototype of an  $n=1$  system.

In order to obtain any interesting behavior there must be some interactions between the spins. The standard, simplest Hamiltonian for an Ising model, given that there are  $N$  lattice sites, is

$$\mathcal{H}_N(\{s_i\}) = -H \sum_{i=1}^N s_i - \sum_{\langle i,j \rangle} J s_i \cdot s_j. \quad (4.2)$$

The first term in  $\mathcal{H}_N$  takes account of any externally applied magnetic field,  $H$ . This term on its own would give us only a paramagnet. The second term describes the interactions between spins. For  $J > 0$  it is of ferromagnetic character (approximating the so-called exchange coupling) and tends to line up the neighboring spins,  $s_i$  and  $s_j$ , in the same direction. The notation  $\langle i,j \rangle$  indicates that the sum extends only over nearest neighbor pairs of lattice sites. (Sometimes the notation  $[i,j]$  is used.) When it is appropriate to consider interactions of longer range, the sum must run over all pairs  $(i,j)$  and the coupling or "exchange parameter",  $J$ , is replaced by  $J_{ij} = J(\vec{r}_i, \vec{r}_j)$  with  $J_{ij} \equiv J(\vec{r}_j - \vec{r}_i)$  in the normal, translationally invariant case.

### 4.3 Solution of the one-dimensional Ising model

The one-dimensional case of this simple model was originally solved by E. Ising in his 1925 thesis: the solution can be found in a number of text books on statistical mechanics. It will not be repeated here in full but some of the essential steps will be outlined. (They should be part of any graduate level course in statistical mechanics!) The modern way of solving the Ising model in one-dimension is first to recognize that what has to be calculated is the partition function

$$Z_N[\mathcal{H}] = \text{Tr}_N \{ e^{-\mathcal{H}(s)/k_B T} \} = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} e^{-\mathcal{H}(s_1, \dots, s_N)/k_B T}. \quad (4.3)$$

Because there are so many spins the calculation is difficult. On facing a hostile army in overwhelming numbers the classical tacticians advised: "divide and conquer", or if possible, "pick them off one by one". The easy method of solving the one-dimensional Ising model (which, incidentally, is not the method Ising himself used) follows the second adage. One considers a linear lattice of  $N$  spins in which the spin summations,  $\sum_{s_i=\pm 1}$ , have been done over all spins  $s_i$  ( $i=1,2,\dots$ ) except the last one,  $s_N$ . Then one asks what happens on adding one further spin,  $s_{N+1}$ . Assuming one knows the "partial partition function" for the  $N$ -spin system, say  $Z_N[\mathcal{H}; s_N]$ , one sees that only one more summation, i.e., over  $s_N$ , is needed to compute  $Z_{N+1}[\mathcal{H}; s_{N+1}]$ , and so on. Finally one must take the thermodynamic limit in order to compute the bulk free energy density and see if it has any singularities that might represent phase transitions or critical behavior. Explicitly, we define the reduced free energy density via

$$f(T, H) = \lim_{N \rightarrow \infty} (N v_0)^{-1} \ln Z_N[\mathcal{H}], \quad (4.4)$$

where  $v_0$  is the volume (length in this  $d=1$  case) per site. It turns out that the process of sequentially adding one spin at a time can be done very simply and directly in terms of a  $2 \times 2$  matrix, which depends on  $T$  and  $H$ . Furthermore, on taking the thermodynamic limit one finds that for any boundary conditions all one needs to know about the matrix is its largest eigenvalue, say,  $\Lambda_{\max}(T, H)$ . In terms of this maximal eigenvalue (which, since the matrix is nonnegative, has to be real) one simply has

$$f(T, H) = \ln \Lambda_{\max}(T, H). \quad (4.5)$$

On deriving the appropriate  $2 \times 2$  matrix and solving a quadratic equation for its eigenvalues the answer that comes out (after subtracting the harmless ground state contribution) is

$$f(T, H) = \ln \left[ \cosh h + \sqrt{\sinh^2 h + x} \right]. \quad (4.6)$$

Here we have introduced the reduced variables

$$h = H/k_B T, \text{ and } K = J/k_B T, \quad (4.7)$$

and taken

$$x = e^{-4K} = \exp(-4J/k_B T). \quad (4.8)$$

The variable  $x$  is a temperature-like quantity which vanishes when  $T \rightarrow 0$ . For  $H = 0$  and  $T \rightarrow 0$  the ground state will be attained and, clearly, this will correspond to all the spins pointing in the same direction, either 'up' or 'down'. To reverse a particular spin then requires an energy input of  $4J$  because of the interaction of this spin with its two neighbors. So this first, 'single-spin-flip' excited state comes with a Boltzmann factor of  $\exp(-4J/k_B T)$  which demonstrates why, in Ising systems,  $x$  is the natural low temperature in terms of which one would, a priori expect simple, analytic behavior.

What, historically, was disappointing to the early investigators was that this model seemed not to give any phase transition whatsoever! In particular it displays no sharp specific heat anomaly at any finite (nonzero) temperature. This can be seen by setting  $H=0$  in (4.6) so obtaining  $f = \ln(1+x^{1/2})$  which is clearly a smooth function of  $x$  or  $T$  right down to the absolute zero. In fact it is now known that one-dimensional systems with quite general finite range pair interactions cannot have phase transitions at any nonzero temperature. Nevertheless, even the simplest nearest neighbor, one-dimensional Ising model does have a transition at  $T = 0$  which can properly be regarded as a critical point! This can be seen, for example, by studying small  $x$  (or  $T \rightarrow 0$ ) for which one has

$$f = \ln(1+x^{1/2}) = x^{1/2} [1 + O(x^{1/2})]. \quad (4.9)$$

Since a priori we would have expected  $f$  to vary simply as  $x$  (as explained above) we see that the power  $1/2$  must really be regarded as a special sort of critical exponent (for the exponential temperature variable  $x$ ).

At first sight it may seem a bit artificial to regard this effect as signifying a phase transition but, in fact, the case for doing so is strong! Consider, for example, the magnetization isotherms for the model, as sketched in Fig. 4.3. At  $T = 0$  a discontinuity in  $M$  occurs as  $H$  passes through zero, whereas for any nonzero temperature  $M$  varies smoothly with  $H$ . Moreover, as  $T \rightarrow 0$  the susceptibility,  $\chi = (\partial M / \partial H)_T$  diverges very strongly: in fact, one has  $\chi \sim x^{-\gamma_x}$  where  $\gamma_x = 1/2$ . This exponentially strong divergence of  $\chi$  with  $T \rightarrow 0$  should be contrasted with the simple paramagnetic behavior  $\chi \sim 1/T$ . Similar exponents can be defined for all other quantities of interest. Also scaling theory

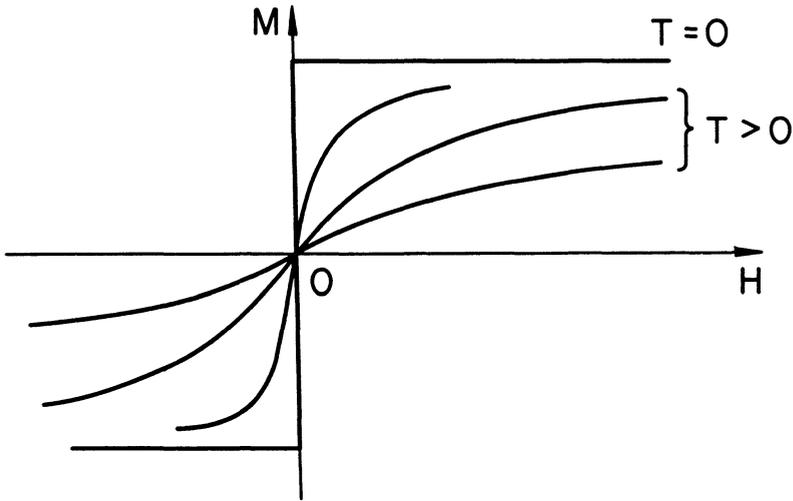


Fig. 4.3 Magnetization curves for the one-dimensional Ising model.

can be applied with  $(T-T_c)$  replaced by  $x$ . (It is a good exercise for the student to check this and find the scaling functions!)

#### 4.4 The two-dimensional Ising model

We will, here, just present some of the exact results, calculated for the two-dimensional Ising model, which reinforce our belief in the scaling hypothesis. This model is almost synonymous with the name of Lars Onsager who solved it analytically in 1944 by a generalization of the matrix method sketched above for the one-dimensional model. The first important result to emerge was that there was indeed a phase transition at  $T_c > 0$ . At the critical point the specific heat diverged with exponents

$$\alpha = \alpha' = 0 \quad (\log), \quad (d=2, \text{ Ising}). \quad (4.10)$$

The logarithmic divergence was the first striking demonstration that the classical theory was quite wrong! Onsager also showed that

$$\beta = \frac{1}{8}, \quad (d=2, \text{ Ising}), \quad (4.11)$$

which is very different from the classical value of  $1/2$ . (Onsager announced his result at an early stage but delayed publishing his derivation. The first published calculation for  $\beta$  is due to C. N. Yang.) Onsager also calculated the correlation length from which one finds

$$\nu = \nu' = 1, \quad (d=2, \text{ Ising}), \quad (4.12)$$

which contrasts with the classical value  $1/2$ . Finally he set up the calculations which lead to the demonstration that

$$\eta = \frac{1}{4}, \quad (d=2, \text{ Ising}), \quad (4.13)$$

in disagreement with the classical Ornstein-Zernike prediction  $\eta=0$ . Later investigators showed that

$$\gamma = \gamma' = 1 \frac{3}{4} \quad \text{and} \quad \delta = 15, \quad (d=2, \text{ Ising}), \quad (4.14)$$

the classical values being  $\gamma = \gamma' = 1$  and  $\delta = 3$ .

Note that precisely the same exponent values apply to all soluble two-dimensional Ising models, which means all those with only nearest-neighbor interactions. Furthermore, the exact values satisfy all the scaling and all the hyperscaling relations derived earlier. Unfortunately, the full scaling of the equation of state itself cannot be checked because the model has not yet been solved in the presence of a magnetic field. However, scaling of the pair correlation functions can be checked in detail. (One might remark that the exact solution of the two-dimensional Ising model in a magnetic field would probably teach us much more about phase transitions and critical phenomena, at this stage, than the exact solution of the three-dimensional model in zero magnetic field: in particular, it would reveal the nature of the singularities on the approach to a first order transition, a fascinating but subtle matter beyond the scope of these lectures.)

#### 4.5 Ising model in three dimensions: series expansions

One thing that has been very clearly revealed by the exact analysis of two-dimensional Ising models and by comparison of the results with experiment and with the classical predictions, is that the dimensionality,  $d$ , must play a crucial role in determining the critical exponents. It is obvious, therefore, that one should also want to study the three-dimensional Ising model! This model cannot be solved analytically in the same way that Onsager solved the two-dimensional one (although there were, initially, quite a few attempts). However, answers to many of the crucial questions have been obtained to rather good precision by means of numerical

"solutions". The method that has been used is the technique of exact series expansions (pioneered by Cyril Domb<sup>13</sup>). To see how this works, the easiest quantity to consider is the reduced susceptibility

$$\chi = \chi_T / \chi_T^{\text{ideal}}. \quad (4.15)$$

At high temperatures, as  $T \rightarrow \infty$ , any spin system will behave more and more like an ideal paramagnet, and so  $\chi \rightarrow 1$ . Thus we may seek an approximation which will approach this result in the high temperature limit but which we will, in fact, attempt to use also near in the critical region. To this end, recall that for  $H = 0$  the partition function is given by

$$Z_N = \text{Tr}_N \{ e^{-\mathcal{H}/k_B T} \} = \sum_{\{s_i = \pm 1\}} \exp \left[ K \sum_{\langle i, j \rangle} s_i s_j \right]. \quad (4.16)$$

It is a function of the single parameter  $K = -J/k_B T$  which becomes small when  $T \rightarrow \infty$ . It is thus natural to search for an expansion of the properties of the Ising model, for arbitrary dimensionalities, in powers of K. The most direct approach is to make use of the identity (see also the more detailed discussion in Appendix C),

$$e^x = 1 + x + \frac{1}{2} x^2 + \dots, \quad (4.17)$$

and thence obtain an expansion for  $Z_N$  in powers of  $K$ . However, to simplify the calculations it turns out to be better to introduce another temperature-like variable, namely,

$$v = \tanh K = K + O(K^3) \rightarrow 0 \text{ as } T \rightarrow \infty. \quad (4.18)$$

Then, and this is not very difficult to show for the first few terms, one finds, for example, that the expansion for  $\chi$  in powers of  $v$  for the simple cubic Ising lattice is

$$\chi = 1 + 6v + 30v^2 + 150v^3 + 726v^4 + 3510v^5 + 16710v^6 + 79494v^7 + 375174v^8 + 1769686v^9 + \dots + 86228667894v^{16} + 401225391222v^{17} + \dots, \quad (4.19)$$

where we now know even the coefficient of  $v^{19}$  although it is too long to write here! The coefficients have a fairly simple interpretation. Starting at the origin site on a simple cubic lattice,  $a_1=6$  is the number of ways to reach the first nearest neighbor sites;  $a_2=30$  to reach the second nearest neighbors and so on, except that from  $a_5=3510$  onwards further complications enter. The general coefficient,  $a_m$ , is, in first approximation, just the total number of distinct self-avoiding walks of  $m$  steps starting from the origin: however, this has to be corrected by allowing for a "gas of polygons" that use some of the  $m$  available

lattice bonds. A great deal of effort has been expended in calculating the higher order coefficients; computers are of some assistance but they by no means make the task trivial: indeed for many years they could not compete effectively with systematic hand calculations!

Now it is quite evident that the coefficients  $a_m$  in (4.19) are increasing in a rather regular fashion with  $m$ . To study this, let us examine the ratios  $\mu_m = a_m/a_{m-1}$  of successive coefficients by plotting them versus  $1/m$ . Readers are urged to do this for themselves using the data given in (4.19). (Note that the ratios for  $m = 1$  to 9 and for  $m = 17$  can be plotted). Those lacking the energy or time may consult the literature.<sup>14</sup> For lattices like the triangular or fcc lattices that contain three-sided polygons, i.e., triangles, one finds that successive ratios fall close to a straight line of positive slope vs  $1/m$ . For 'loose-packed' lattices like the square and simple cubic lattice (containing no triangles) there is an odd-even alternation of the ratios but both sets of ratios rapidly approach a similar straight line!

What does this mean? Certainly one may conclude that, to apparently ever better approximation as  $m$  increases, one may write

$$\mu_m = \frac{a_m}{a_{m-1}} \equiv \mu_\infty [1 + \frac{g}{m} + \dots], \quad (4.20)$$

where the dots stand for terms vanishing more rapidly than  $1/m$  (see below). Here  $\mu_\infty$  represents the (asymptotic) intercept of the line of ratios with the  $m = \infty$ , i.e.,  $1/m = 0$  axis, while  $g$  represents a dimensionless measure of the slope of the plot. But again, what is this telling us?

Now it is a simple matter to see, with the aid of the binomial expansion, that the power series expansion of the function

$$A_g(v) = \frac{1}{[1-(v/v_c)]^{1+g}} = \sum_{m=0}^{\infty} \binom{-1-g}{m} \left(\frac{v}{v_c}\right)^m = 1 + a_1 v + a_2 v^2 + \dots, \quad (4.21)$$

produces coefficients which generate ratios given exactly by

$$\mu_m = \frac{a_m}{a_{m-1}} = \frac{(g+m)}{mv_c} = \frac{1}{v_c} \left[1 + \frac{g}{m}\right], \quad (4.22)$$

that is, which fall exactly on a straight line in a plot versus  $1/m$ . Further we see that the limiting ratio determines the point of divergence of the series as  $v \rightarrow v_c^-$  via

$$\mu_\infty = 1/v_c. \quad (4.23)$$

More importantly, however, the slope,  $g$ , of the ratio plot evidently tells us the exponent of divergence! (Higher order terms in a ratio plot correspond, of course,

to deviations from the ideal, pure binomial form (4.20): see further below).

To apply this observation to the Ising model note the successive relations

$$\chi(v) \sim [1 - (\frac{v}{v_c})]^{-(1+g)} \sim (K_c - K)^{-(1+g)} \sim 1/(T - T_c)^{1+g}. \quad (4.24)$$

So that, on recalling the exponent definition  $\chi \sim 1/t^\gamma$ , we see that the susceptibility ratio plots provide an estimate of the exponent  $\gamma$  via

$$\gamma = 1 + g. \quad (4.25)$$

In a similar way  $\mu_\infty$ , the intercept at  $1/\mu = 0$ , is directly related to the critical temperature. In fact one finds for the square, honeycomb and triangular lattices that the ratios  $\mu_m$  extrapolate to the exactly known transition temperatures found by Onsager for these lattices to a precision of 1 in  $10^4$  to  $10^5$  or better (depending on the length of the series used). Likewise the estimates of  $\gamma$  for the two-dimensional lattices come out very close (to within  $\pm 0.01$  to  $\pm 0.003$ ) to the exact value  $\gamma = 1.75$ .

For three-dimensional Ising lattices a parallel analysis yields estimates for the critical points (which, of course, depend on the lattice) of, apparently, quite comparable precision. Furthermore, to within the apparent accuracy all three-dimensional lattices studied (sc, bcc, fcc, and diamond) yield the same value of  $\gamma$ . This checks the concept of universality. In 1966, a reasonable best estimate obtained by these methods and various refinements was quoted as  $\gamma = 1.250 \pm 3$ , where the assessed uncertainty refers to the last decimal place given. More recent work, stimulated in particular by renormalization group calculations, by the availability of longer series and by methods for studying corrections to (4.21) (see below) leads to series estimates like<sup>15</sup>

$$\gamma = 1.239 \pm 2. \quad (4.26)$$

which are about 1% lower. These latest estimates agree remarkably well with experimentally measured values for fluids, binary alloys, and other  $n=1$  systems. Other exponent values obtained for three-dimensional Ising models by these techniques are:

$$\alpha \approx 0.105 \pm 10, \nu \approx 0.632 \pm 2, \text{ and } \beta \approx 0.328 \pm 8. \quad (4.27)$$

To within the apparent uncertainties these estimates satisfy all the exponent relations including the hyperscaling relations.

Series extrapolation methods are not applicable only to critical phenomena, but can be used also in many other situations. Notice the fundamental difference in the approach from the normal truncation method of just adding up those finite number of

terms of an infinite series that one has been able to calculate, and merely stopping there! Even for a convergent series this latter method must fail completely in critical phenomena since it can yield no more than a finite polynomial which could never reproduce the divergence to infinity of the susceptibility or other more subtle singular behavior that occurs at critical points.

When a large number of terms in a series have been calculated one can hope to estimate the higher order coefficients in the asymptotic form (4.20) for the ratios. The actual variation of  $\chi$  for a reasonably wide range of  $t$  is, as seen earlier, given by

$$\chi = \frac{C}{t^\gamma} (1 + c_\theta t^\theta + \dots + c_1 t + c_2 t^2 + \dots). \quad (4.28)$$

If the leading correction term is  $c_1 t$  (i.e.,  $\theta > 1$ ), it turns out that  $c_1$  can be related to the coefficient of a  $1/m^2$  term in (4.20). If, however, there is a singular correction term,  $c_\theta t^\theta$ , with  $\theta < 1$ , this will show up as a dominant non-integral power  $1/m^{1+\theta}$  in the expansion of  $\nu_m$ . Extracting reliable information from these terms has proved difficult but, as indicated before (4.26), there has been recent progress on the problem.

In the case of the spontaneous magnetization of three-dimensional Ising models the series turn out to be much more erratic in appearance. For example, the fcc Ising lattice yields, after much labor, the low temperature expansion

$$\begin{aligned} M_0(T) = & 1 - 2x^{12} - 24x^{22} + 26x^{24} + 0+0-48x^{30} - 252x^{32} + 720x^{34} \\ & - 438x^{36} - 192x^{38} - 984x^{40} - 1008x^{42} + 12924x^{44} - 19536x^{46} \\ & + 3062x^{48} - \dots + 400576168x^{78} - 410287368x^{80} + \dots \end{aligned} \quad (4.29)$$

The ratio method evidently fails completely for a case like this! Fortunately, however, there is another method, the so-called Padé approximant technique (propounded originally for this sort of problem by G. A. Baker, Jr. and J. L. Gammel). Padé approximant methods and their generalizations are able to handle such series and yield estimates for  $\beta$  and other low temperature exponents like those quoted above. In addition Padé approximants and their extensions provide efficient methods of approximate summation of series over the whole range of temperatures.

Considerable effort by various research groups has gone into these series expansion methods in an effort to calculate increasingly precise and reliable values for all the critical exponents. An important historical motivation has been the desire to check universality over a variety of models beyond the simplest Ising models. Do the exponents depend on anything besides  $d$  and  $n$ , e.g., lattice structure, quantum effects, the magnitude  $S$  of the spin, further neighbor couplings and so on? As more extensive results have become available these calculations have

increasingly confirmed the surmise that for systems with interactions of finite range  $d$  and  $n$  are the only relevant quantities; other parameters embodying the fundamental constants  $\hbar$ ,  $c$ ,  $e$  etc., apparently play no role at all in this question! This was a great surprise to many of the earlier workers in the field but now, thanks to the work of Kadanoff, Wilson, and others, the reasons for this fact are much better understood. As we shall see, the renormalization group concept provides a natural explanation.

#### 4.6 The n-vector spin models

As pointed out earlier, we need to study not just one model but, rather, a variety of models. A natural hierarchy of classical spin models is represented by the following choices of spin variables:

(a) Ising model:  $s_i = \pm 1 \quad (n=1),$

(b) XY model:  $\vec{s}_i = (s_i^x, s_i^y), \quad |\vec{s}_i| = 1 \quad (n=2);$

(c) Heisenberg model:  $\vec{s}_i = (s_i^x, s_i^y, s_i^z), \quad |\vec{s}_i| = 1 \quad (n=3);$

and finally

(d) n-vector models:  $\vec{s}_i = (s_i^{(1)}, s_i^{(2)}, \dots, s_i^{(n)}), \quad |\vec{s}_i| = 1 \quad (\text{general } n).$

In these models the spin components are simply regarded as classical variables and there are no problems associated with noncommutability, as in the more realistic quantal Heisenberg model with spin  $S < \infty$  (but still with  $n=3$ ). The total spin magnitudes may be normalized to unity as indicated or, as is more appropriate if one wants to consider large  $n$ , to  $|\vec{s}_i|^2 = n$ .

The Hamiltonian in the simplest case would be

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{s}_i \cdot \vec{s}_j - H \sum_i s_i^{(1)}, \quad (4.30)$$

where the first term is the coupling between nearest neighbor spins and the second represents the interaction with the external magnetic field which one supposes is applied in such a way that it couples only to the first component of the vectors  $\vec{s}_i$ . With the spin normalization set by  $|\vec{s}_i|^2 = n$ , the free energy density of interest is derived from the Hamiltonian via

$$f[\mathcal{K}] = \lim_{N \rightarrow \infty} -F_N / nNa^d k_B T = \lim_{N \rightarrow \infty} \frac{1}{nNa^d} \ln Z_N[\mathcal{K}], \quad (4.31)$$

where we have expressed the cell volume  $v_0$  in terms of the lattice spacing a supposing, for simplicity, a  $d$ -dimensional hypercubic lattice. The factor of  $n$  in the denominator means that  $f$  is a free energy evaluated not only per unit volume, but also per single spin component. When defined in this way one can extract sensible results even when (as proposed by H. E. Stanley) the limit  $n \rightarrow \infty$  i.e., of an infinite number of spin components, is taken! It transpires that this leads to the so-called spherical model invented by Mark Kac. The interest in this seemingly most artificial limit is not because the model is at all physical, but rather because it can be solved exactly and it embodies important features, characteristic of models with  $n > 1$  which cannot be studied in the scalar case of Ising models. The exponent values that emerge for the exact solutions are

$$\eta = 0, \text{ and } \beta = \frac{1}{2}, \text{ for all } d, \quad (4.32)$$

and, with a striking dimensional-dependence,

$$\begin{aligned} \gamma &= 2\nu = 2/(d-2), \text{ for } d < 4, \\ &= 1, \text{ for } d > 4, \end{aligned} \quad (4.33)$$

and, for the specific heat:

$$\alpha = \epsilon/(d-2) \text{ where } \epsilon = 4-d. \quad (4.34)$$

We see here the appearance of the dimensionality parameter  $\epsilon = 4-d$ ; later on we will use this as a crucial expansion parameter in renormalization group theory. From a mathematical point of view it makes perfectly good sense (with a little care) to treat the spatial dimensionality,  $d$ , as a continuous variable even though it is only integral values that have a direct physical meaning. In this model there are clearly two special values of  $d$ , namely,  $d = 4$ , called the upper borderline (or marginal) dimensionality; and  $d = 2$ , called the lower borderline dimensionality. For  $d > 4$ , classical theory is seen to work. At  $d = 2$  the critical temperature vanishes,  $T_c = 0$ , and for  $d < 2$  there is no phase transition at all. As  $d \rightarrow 2+$  some of the critical exponents diverge to  $\infty$  as can be seen from the results above. It turns out that these two borderlines apply for all  $n > 1$ . Symptomatic of another feature that appears generally at borderline dimensionalities is that for  $d = 4$  one finds that the full critical behavior of the susceptibility is

$$\chi \approx \frac{C}{t} \ln t \text{ as } t \rightarrow 0+, \quad (4.35)$$

i.e., there is a logarithmic correction factor to the leading power law behavior.

The appearance of a logarithmic factor at the borderline dimensionality  $d = 4$  suggests, correctly, that simple power law scaling forms must break down somewhat at this margin. However, for  $d < 4$  (including continuous values of  $d$ ) the spherical model satisfies all the exponent relations and the thermodynamic and correlation functions scale completely in the standard way. Above  $d = 4$  one finds that scaling again works: indeed the classical theory is, asymptotically, valid. However, that does mean that the hyperscaling relations (that involve  $d$  explicitly) must fail for  $d > 4$ . This observation indicates the somewhat different status that should be accorded the hyperscaling relations: below  $d = 4$ , however, they are precisely satisfied in the spherical model.

While the physically realizable values of  $d$  are severely limited, the values of  $n$  are much less restricted. Thus, as we have seen  $n = 1, 2$  and  $3$  are commonplace; but the case  $n = 18$ , for example, is of interest in describing the superfluid properties of  $\text{He}^3$  (where the order parameter can be represented as a complex  $3 \times 3$  matrix) and values as high as  $n = 48$  are at least conceivable in connection with certain incommensurate phase transitions. In the opposite direction,  $n = 0$  turns out to describe the pure self-avoiding walk problem or, in physical terms, the excluded volume problem for polymers in solution. Even a negative number of components may be considered! Thus the case  $n = -2$  has certain attractive analytical features.<sup>16</sup>

#### 4.7 Continuous spin models

In these models the spin is again regarded as an  $n$ -component classical vector, but now each component is allowed to range from  $+\infty$  to  $-\infty$ , so we have

$$\vec{s}_i = (s_i^{(1)}, \dots, s_i^{(n)}), \text{ with } -\infty < s_i^{(\mu)} < +\infty. \quad (4.36)$$

As regards the interactions between the spins, the Hamiltonian is just the same as before, that is

$$\mathcal{H}_{\text{int}} = -J \sum_{\langle i,j \rangle} \vec{s}_i \cdot \vec{s}_j - H \sum_i s_i^{(1)}. \quad (4.37)$$

However, there is now a further feature one must consider: this is the spin distribution. If one were not to place some sort of constraint on the magnitudes of the spins,  $\vec{s}_i$ , or on the way in which the components  $s_i^{(\mu)}$  can be distributed then the total energy could be made indefinitely large and negative over an infinitely large region of phase space; the partition function would thus diverge while all the spins become infinitely large! We may, however, choose to regard the standard Ising

model as a special class of continuous spin models, and this provides us with some guidance as to what to do to obtain a sensible model. An Ising spin  $s_i$  can be considered as a continuous, one-component spin, but with the constraint that it can take on only the values  $\pm 1$ . Another way of describing this would be to say that each spin is subject to a distribution function or spin weighting function

$$e^{-w(s_i)} = \delta(s_i + 1) + \delta(s_i - 1), \quad (\text{all } i). \quad (4.38)$$

With the use of this weight function, the trace sums that are involved in calculating the partition function of an Ising spin system can be transformed into integrals so that we obtain

$$Z_N[\mathcal{H}] = \int_{-\infty}^{\infty} ds_1 \dots \int_{-\infty}^{\infty} ds_N e^{-\mathcal{H}_{\text{int}}/k_B T} \prod_{i=1}^N e^{-w(s_i)}. \quad (4.39)$$

(In the case of the n-vector model with  $n > 1$  each integral becomes an n-fold integral over the components  $s_i^{(\mu)}$ .)

The simplest generalization of (4.37) that provides a genuinely continuous spin distribution, is the Gaussian model. This model (also due to Kac) is obtained by setting

$$w(\vec{s}_i) = \frac{1}{2} |\vec{s}_i|^2, \quad (\text{all } i). \quad (4.40)$$

The integrand in (4.38) is now just an exponential of a quadratic expression. Consequently the calculation of  $Z_N[\mathcal{H}]$  reduces, after diagonalizing the quadratic form, simply to taking a product of Gaussian integrals. The model is thus exactly soluble! The Gaussian model happens to correspond precisely to the artificial limit  $n = -2$  mentioned above! The exponent values that emerge from its solution are

$$\eta = 0, \text{ and } \gamma = 2\nu = 1 \text{ for all } d, \quad (4.41)$$

while the specific heat exponent is given by

$$\alpha = \begin{cases} \frac{1}{2} \varepsilon & \text{for } d < 4, \\ 0 & \text{for } d > 4. \end{cases} \quad (4.42)$$

Unfortunately, the Gaussian model has serious shortcomings. Its worst feature is that it has no low temperature behavior! The reason for this is that the exponential decrease  $e^{-|\vec{s}|^2}$  of the spin weighting function for large  $|\vec{s}|$  is just not rapid enough to keep the attractive coupling terms under control when  $T$  is too small: as a consequence the integrals diverge and the model collapses!

To overcome this fatal defect we introduce a generalization of fundamental

significance: this is the so-called  $s^4$  model now often called the LGW or Landau-Ginzburg-Wilson model. In this model the spin weighting function is taken as

$$e^{-w(\vec{s})} = e^{-\frac{1}{2}|\vec{s}|^2 - \tilde{u}|\vec{s}|^4} \quad \text{with } \tilde{u} > 0. \quad (4.43)$$

The effect of the  $\tilde{u}|\vec{s}|^4$  term is to pull the tails of the Gaussian weighting function down rapidly and hence give it a squarer looking shape (see Fig. 4.4). It then approximates the Ising model more closely, at least insofar as there is little weight for large  $s$ ; unphysical features of the Gaussian model below  $T_c$  are quite absent. It is widely believed that the exponent values for the scalar ( $n = 1$ )  $s^4$

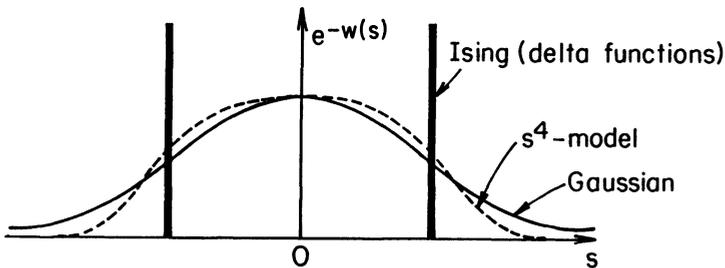


Fig. 4.4 Schematic comparison of the spin weighting functions  $\exp[-w(s)]$  for the Ising, Gaussian and  $s^4$  models.

and Ising models are exactly the same. The only sad result of including the  $s^4$  terms in the exponential is that the integrals defining the partition function in (4.37) can now no longer be done exactly as before! This, however, is where  $\epsilon$ -expansions, which will be discussed in a later section, have a valuable role to play.

The weighting function contributions,  $w(s_i)$ , are often treated as an integral part of the overall Hamiltonian: thus one writes the total reduced Hamiltonian as

$$\bar{\mathcal{H}}(s) = -\mathcal{H}_{\text{int}}(\vec{s}_1, \dots, \vec{s}_N)/k_B T - \sum_{i=1}^N w(\vec{s}_i), \quad (4.44)$$

and the partition function is given by

$$Z_N[\mathcal{H}] = \text{Tr}_N \{ e^{\bar{\mathcal{H}}(\vec{s}_1, \dots, \vec{s}_N)} \} = \int_{-\infty}^{\infty} d^n s_1 \dots \int_{-\infty}^{\infty} d^n s_N e^{\bar{\mathcal{H}}(\vec{s}_1, \dots, \vec{s}_N)}. \quad (4.45)$$

These expressions will be our starting point in discussing the renormalization group  $\epsilon$  expansions.

At first sight the continuous spin models with smooth weighting functions seem intrinsically different from the discrete spin or fixed-length spin models. However, as we show in Appendix A, they can in fact represent the discrete and fixed length models exactly!

## 5 Renormalization Group Theory

### 5.1 Preamble

To start with let us concentrate on the essence of renormalization group theory, putting the ideas in their simplest form. An analogy may be useful to give some perspective. In the progression from classical mechanics to a full account of quantum mechanics one starts first of all with the Bohr-Sommerfeld model or picture. Although this represents only a crude approximation, it nevertheless introduces some important ideas, such as quantization and energy levels, and it provides an explanation for the existence of discrete spectral lines and other specifically quantum-mechanical phenomena. Naturally one wants to move on from there to Schrodinger's equation and the particle-wave duality, to Bose-Einstein and Fermi-Dirac statistics, to Dirac's equation, and to quantized field theory! Nevertheless, it is instructive to start with the simplest embodiment of the most basic ideas.

In critical phenomena, the counterpart of quantization is the concept of a renormalization group transformation. The simplest such transformation which corresponds to the Bohr-Sommerfeld picture, is realized in the renormalization group treatment of the one dimensional nearest neighbor Ising model. This model can, as we have seen, be solved exactly in a fairly easy way but an analysis using a renormalization group approach still serves to introduce some important concepts. From there one hopes to progress to more subtle models. In general, the most basic task of renormalization group theory is to explain scaling, to show us where the critical exponents come from, and to explain universality. Beyond that one would like to calculate, more-or-less explicitly, critical exponents and scaling functions. Further, the theory should tell us where the simplest scaling ideas fail and what should replace them when they do!

### 5.2 A renormalization group for the one-dimensional Ising model

What always enters into the partition function, as discussed previously, is the quantity  $-\mathcal{K}/k_B T$ , which for brevity will be called  $\bar{\mathcal{K}}$ . For the one-dimensional nearest neighbor Ising model we therefore have

$$\bar{\mathcal{H}} = -\mathcal{H}/k_B T = K \sum_j s_j s_{j+1} + h \sum_j s_j + C \sum_j 1, \quad (5.1)$$

in which

$$K = J/k_B T, \quad h = H/k_B T, \quad (5.2)$$

while an extra term, equal to  $CN$ , has been added. This has no physical consequences but turns out to serve a useful mathematical purpose in any full renormalization group treatment. It might be regarded as equal to  $-E(0)/k_B T$  where  $E(0)$  is some reference energy.

This "reduced" Hamiltonian,  $\bar{\mathcal{H}}$ , is evidently "equivalent" to the set of variables  $(K, h, C)$ : knowledge of these three variables specifies  $\bar{\mathcal{H}}$ , and therefore determines the free energy completely in the thermodynamic limit. Thus  $\bar{\mathcal{H}}$  can be regarded as a point in the space of the three parameters  $K, h$ , and  $C$ . As the physical variables  $T$  and  $H$  are changed this point moves around.

One of the first approaches to renormalization group theory is to regard it merely as a special way of calculating the partition function which, in this case may be written

$$Z_N[\bar{\mathcal{H}}] = \text{Tr}_N^s \{ e^{\bar{\mathcal{H}}} \} = \frac{1}{2^N} \sum_{S_1=\pm 1} \dots \sum_{S_N=\pm 1} e^{\bar{\mathcal{H}}(s)}. \quad (5.3)$$

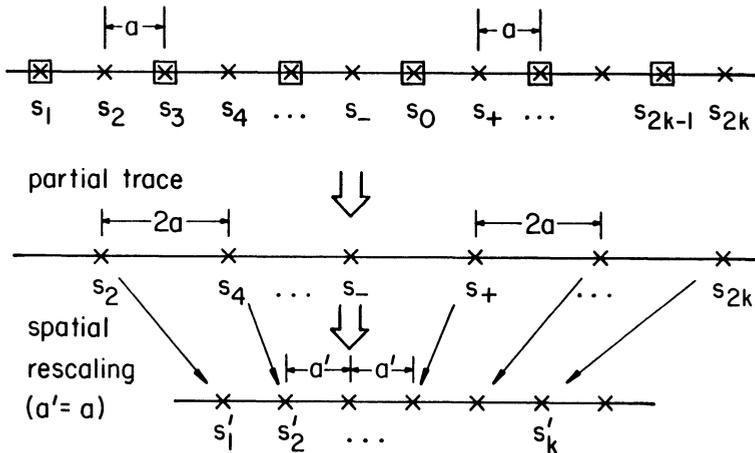


Fig. 5.1 Schematic representation of the simplest "decimation" or "dedecaration" renormalization group for a one-dimensional nearest neighbor Ising model in which a partial trace is taken by summing over alternate spins (boxed) to reduce the number of spins, followed by a spatial rescaling to restore the original appearance of the problem. (see text).

The factor  $1/2^N$  has been introduced here so that  $Z_N$  is conveniently normalized to unity when  $T \rightarrow \infty$ . This too has no physical consequences. The free energy per spin is then

$$f[\bar{\mathcal{H}}] \equiv f(K, h, C) = \frac{1}{N} \ln Z_N[\bar{\mathcal{H}}]. \quad (5.4)$$

The renormalization group method of tackling the problem of evaluating  $f[\bar{\mathcal{H}}]$  is, like the matrix method discussed earlier, one of "divide and conquer". The idea is that instead of trying to do all the  $N$  spin summations at once, one should somehow do the summations over only some of the spins at one time, in such a way as to try to preserve the system looking as much as possible like it did before the summation, and in such a way that a spatial rescaling of the system is effected.

To achieve these ends in the present case we perform a partial trace by summing over only every second spin variable along the chain, leaving the alternate spins unaffected.<sup>17</sup> In this way we obtain a "renormalized" chain with only half the number of original spins as illustrated schematically in Fig. 5.1. To see what this really entails we first of all write the total Boltzmann weight in the factored form

$$\begin{aligned} e^{-\bar{\mathcal{H}}} &= \dots e^{Ks_-s_0 + \frac{1}{2}h(s_- + s_0) + C} e^{Ks_0s_+ + \frac{1}{2}h(s_0 + s_+) + C} \dots \\ &= \dots \rho(s_-, s_0) \rho(s_0, s_+) \dots, \end{aligned} \quad (5.5)$$

where  $\rho$  can be regarded as the Boltzmann factor for a nearest neighbor "bond", and depends only on the two spins lying at the ends of that bond. The spins  $s_0$ ,  $s_-$ , and  $s_+$  just denote one of the typical spins over which we wish to sum, together with its two nearest neighbors, respectively. The partial trace to be taken will eliminate the spin variable  $s_0$ , and result in a new Boltzmann factor, namely,

$$\rho'(s_-, s_+) = \frac{1}{2} \sum_{s_0=\pm 1} \rho(s_-, s_0) \rho(s_0, s_+), \quad (5.6)$$

for the new "bond" connecting  $s_-$  and  $s_+$ . It is unlikely that the new Boltzmann factor will look exactly like the old one and so it has been written with a prime and is said to be "renormalized". The factor  $\frac{1}{2}$  is included in this relation because with each spin eliminated one must remove a factor of  $\frac{1}{2}$  from the overall normalizing factor for the partition function in (5.3). This process of eliminating spins is usually called "decimation" although "secundation" might be a more appropriate term in view of the fact it is every second spin that is "killed off" rather than every tenth one (from which Roman disciplinary procedure the word derives!) The term "dedecoration" is sometimes also used since the process is the reverse of "decorating" every bond with a new spin.

Now the renormalization group ideal is to be able to express the new bond factor,  $\rho'$ , in the same basic form as the old one i.e., we would like to have

$$\rho'(s_-, s_+) = e^{K's_-s_+ + \frac{1}{2}h'(s_- + s_+) + C'}, \quad (5.7)$$

so that the new spin chain would also be completely Ising-like. There is no reason, however, to expect that the new or renormalized parameters,  $K'$ ,  $h'$  and  $C'$  should take the same values as the old ones. Rather we suppose the new parameters define the renormalized Hamiltonian

$$\bar{\mathcal{H}}' \equiv (K', h', C'), \quad (5.8)$$

which will have only half as many spins. If one can indeed achieve this, one is said to have accomplished one step of a renormalization transformation. The result is written formally as

$$\bar{\mathcal{H}}' = \mathcal{R}_b[\bar{\mathcal{H}}]. \quad (5.9)$$

The important parameter  $b$  is called the spatial rescaling factor which in this case is simply equal to 2 (see Fig. 5.1). The change in the number of spins is described by  $b$  since  $N \Rightarrow N' = N/2 = N/b$ . In two spatial dimensions, however, one could consider decimation by knocking out alternate rows of spins and alternate columns of spins: then one would have  $N \Rightarrow N' = N/4 = N/b^2$ . Generally, in a  $d$ -dimensional system the spatial rescaling factor is related to the reduction in the number of degrees of freedom, here simply spins, by

$$N \Rightarrow N' = N/b^d. \quad (5.10)$$

Back in one dimension, we have expressed the hope that  $\rho'(s_-, s_+)$  as obtained from (5.5) can somehow be expressed in the desired form (5.6). Now we have three variables  $K'$ ,  $h'$  and  $C'$  that can be adjusted in order to make this hoped-for identity true. Since  $s_-$  and  $s_+$  can only take on the values  $(+1,+1)$ ,  $(-1,-1)$ ,  $(+1,-1)$  and  $(-1,+1)$ , imposing the identity leads to four matching equations, the last two of which turn out to be identical (because the two ends of a bond are symmetrically related). It is thus an elementary exercise to show that these matching conditions are solved by

$$e^{4K'} = \frac{\cosh(2K + h) \cosh(2K - h)}{\cosh^2 h}, \quad (5.11)$$

$$e^{2h'} = e^{2h} \frac{\cosh(2K + h)}{\cosh(2K - h)}, \quad (5.12)$$

and, lastly, demonstrating why it was important to introduce the "constant" term C,

$$e^{4C'} = e^{8C} \cosh(2K + h) \cosh(2K - h) \cosh^2 h. \quad (5.13)$$

We thus see that the proposed renormalization group transformation can be performed explicitly and exactly in this case. It has resulted in a new Hamiltonian but one retaining the same general form and it entails a reduction in the number of spins and, as we shall see shortly, an associated spatial rescaling.

### 5.2.1 Spatial rescaling and spin correlations

The original Ising model consisted of a chain of spins each separated from its neighbors by the lattice spacing, a. After eliminating every second spin the remaining spins are now a distance 2a apart (see Fig. 5.1). In an effort to have the renormalized model look as much like the old as possible, we rescale all lengths in such a way that the new lattice spacing, a', equals the old one. Under this scale transformation any distance R in the original lattice becomes  $R' = \frac{1}{2} R$  in the new lattice when measured in units of the lattice spacing. In general we have

$$R \Rightarrow R' = R/b. \quad (5.14)$$

This spatial rescaling is of particular importance in relation to the spin-spin correlation function,  $\langle s_0 s_R \rangle$ . First let us notice that it obviously makes sense to renumber the remaining spins so that their labels again run consecutively. Thus, as shown in Fig. 5.1, we take

$$s_2 \Rightarrow s_1', s_4 \Rightarrow s_2', \dots, s_{2k} \Rightarrow s_k', \dots, \quad (5.15)$$

which, if we regard the labels as distance coordinates, is the same as making the identification  $s_{R'} = s_{2R}$ . Second, note that (for this renormalization transformation) since the undecimated spins retain their characters and relation to one another the renormalized correlation function  $\langle s_0' s_{R'} \rangle$  is actually equal to the original correlation function  $\langle s_0 s_{2R} \rangle$ . It follows that if the renormalized correlation length is  $\xi' \equiv \xi[\bar{\mathcal{K}}']$  the original correlation length,  $\xi \equiv \xi[\bar{\mathcal{K}}]$ , is just twice as long! More generally for a spatial rescaling factor b we have derived the important renormalization relation

$$\xi[\bar{\mathcal{K}}] = b\xi[\bar{\mathcal{K}}']. \quad (5.16)$$

We see from this that the renormalization group procedure has the effect of shrinking the correlation length. Hence if we recall the central fact of critical

phenomena, namely, that  $\xi$  becomes indefinitely large as  $t \rightarrow 0$ , we see that a renormalization transformation has the effect of driving a system away from criticality. It transpires that this is, perhaps, the most crucial feature of the method, the one that enables us to focus on critical points!

### 5.2.2 Unitarity

Another crucial aspect that a renormalization transformation should embody may be called unitarity (although the term is not here being used in the sense familiar in matrix theory or quantum mechanics). Recall that the renormalized Hamiltonian arose in the process of carrying out a partial trace over some of the original spin degrees of freedom. If one now simply completes the trace operation by summing over the remaining spins as coupled through the renormalized Hamiltonian to obtain the renormalized partition function, the end result must be the same as if one had performed the entire trace operation in one go. In mathematical terms, we have, first, the partial trace

$$e^{\bar{\mathcal{H}}'(s')} = \text{Tr}_{N', \{e^{\bar{\mathcal{H}}(s)}\}}, \quad (5.17)$$

where  $s'$  stands for the  $N' = N - N'$  spins over which the decimation trace was taken, and then we compute

$$\begin{aligned} Z_N[\bar{\mathcal{H}}'] &\equiv \text{Tr}_{N', \{e^{\bar{\mathcal{H}}'(s')}\}} \\ &= \text{Tr}_{N', \{\text{Tr}_{N', \{e^{\bar{\mathcal{H}}(s)}\}}\}} = \text{Tr}_N[e^{\bar{\mathcal{H}}(s)}] \\ &= Z_N[\bar{\mathcal{H}}]. \end{aligned} \quad (5.18)$$

In other words the partition function is preserved under renormalization or, equivalently, renormalized by the simple factor unity!

This central result yields the law of renormalization for the free energy itself as follows:

$$\begin{aligned} f[\bar{\mathcal{H}}] &= \frac{1}{N} \ln Z_N[\bar{\mathcal{H}}] \\ &= \left(\frac{N'}{N}\right) \frac{1}{N'} \ln Z_{N'}[\bar{\mathcal{H}}'] = b^{-d} f[\bar{\mathcal{H}}']. \end{aligned} \quad (5.19)$$

Note that for the sake of generality we have used (5.10) which applies for any renormalization group.

Having seen how to construct a renormalization group explicitly (in at least one case!), and having identified a number of important general properties, let us enquire into how it may be used to elucidate the nature of a critical point. To this end, we will leave aside the particular algebraic forms that appear in analysing the one-dimensional Ising model, and focus instead upon the more abstract features which they illustrate.

### 5.3 Flow equations, recursion relations, and fixed points

A renormalization transformation, as we have just seen, fundamentally changes a given problem into a new one, which, however, still contains the same essential information as the original one. If we rewrite (5.19) and (5.17) [which entails (5.11)-(5.13)] we may describe the renormalization procedure by a set of flow equations which describe the motion of a point describing the reduced Hamiltonian,  $\bar{\mathcal{H}}$ , in the appropriate space of parameters, which for our Ising chain are the variables  $K$ ,  $h$ , and  $C$ . First, from (5.19) we obtain

$$f[K, h, C] = b^{-d} f[K', h', C'], \quad (5.20)$$

for the free energy, and then we have

$$K' = \mathcal{R}_K(K, h), \quad (5.21)$$

$$h' = \mathcal{R}_h(K, h), \quad (5.22)$$

and, finally,

$$C' = b^d C + \mathcal{R}_0(K, h), \quad (5.23)$$

for the "coupling constants" or "thermodynamic fields" specifying  $\bar{\mathcal{H}}$ . The last three flow equations are sometimes also called the recursion relations for the coupling constants: of course, for the Ising chain they are just (5.11)-(5.13) written in a more abstract form.

Now since the temperature,  $T$ , is built into the parameter  $K$ , these relations

also imply a flow equation for  $T$ . To explore the significance of this let us, for simplicity, first suppose that the magnetic field on our model vanishes, i.e.,  $H = 0$ , so that this flow equation can be written simply as

$$T \Rightarrow T' = \mathcal{R}(T), \quad (5.24)$$

where  $\mathcal{R}(T)$  is the appropriate function of  $T$ . We will also suppose, without justification at this point, that  $\mathcal{R}(T)$  has the form shown in Fig. 5.2, the important feature being that the plot crosses the line  $T' = T$ . We may call the temperature,  $T^*$ , at the crossing a fixed point because it clearly satisfies the relation

$$T^* = \mathcal{R}(T^*), \quad (5.25)$$

which means that when the recursion relation (5.24) is iterated the temperature  $T$  does not change if its initial value is set at  $T = T^*$  i.e., it remains "fixed" at  $T^*$ . Subsequently we will see that  $T^*$ , in fact, represents the critical point (in this simplified, "Bohr-Sommerfeld" description).

Now if one starts with a temperature  $T_1 > T^*$  then, as is easily seen from Fig. 5.2 one finds that  $T_1' = \mathcal{R}(T_1) > T_1$ , so that if the renormalization transformation is iterated, it drives the temperature,  $T$ , further and further from  $T^*$ . The same thing applies if the starting temperature, say  $T_0$ , is below  $T^*$  as also illustrated in the figure. Consequently we see that  $T^*$  is an unstable fixed point: the temperature always moves away from it under successive renormalizations. Of course, there is, in the figure another fixed point at  $T = 0$  which is stable, but this turns out to be only of limited interest. In practice a similar stable fixed point also occurs at  $T = \infty$  as we might have guessed. These totally stable fixed points are usually referred to as "trivial" fixed points.

As we have seen, the flow equation for the correlation length, is

$$\xi(T) = b\xi(T'). \quad (5.26)$$

At a fixed point therefore we must have

$$\xi(T^*) = b\xi(T^*), \quad (5.27)$$

but since  $b > 1$ , this equation has only two possible solutions, namely,

- (i)  $\xi(T^*) = \infty$ , which evidently characterizes a critical point, and

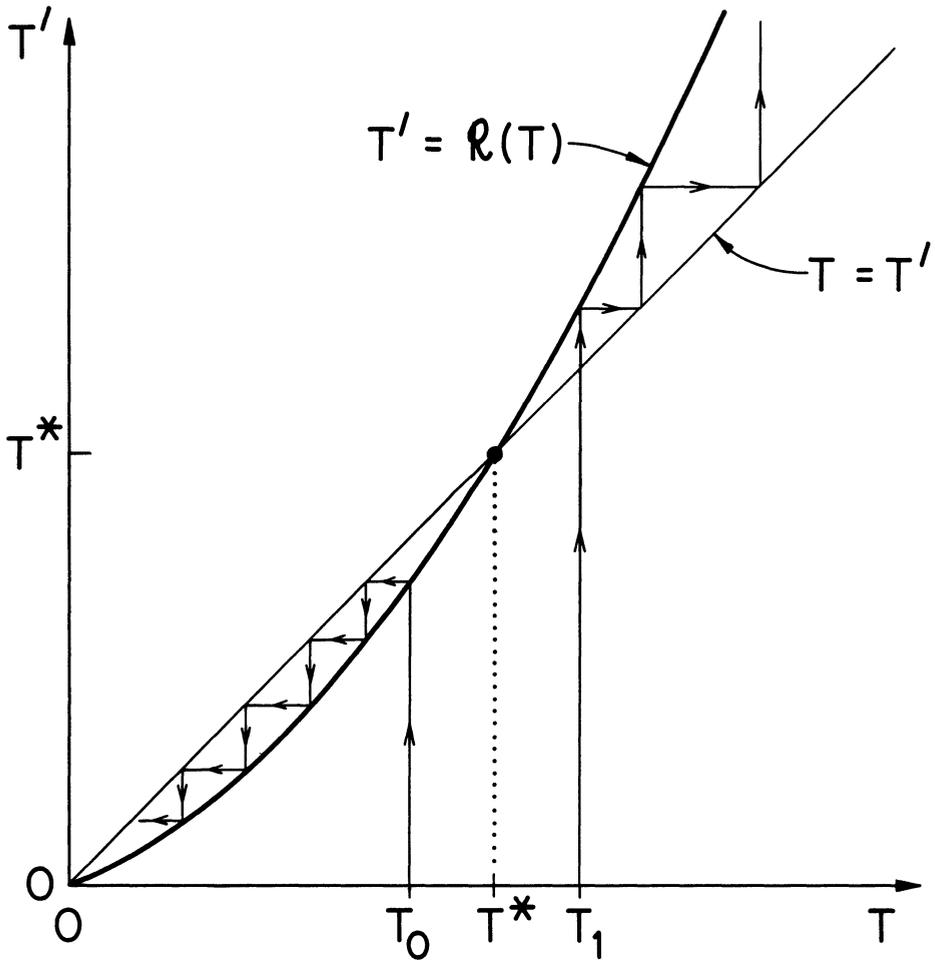


Fig. 5.2 Plot of the renormalization function or recursion relation,  $T' = R(T)$  for the temperature, showing the line  $T' = T$ , the fixed point at  $T = T^*$ , and successive renormalization "flows" resulting from iterating the recursion relation from two starting temperatures  $T_0$  and  $T_1$ .

(ii)  $\xi(T^*) = 0$ , which corresponds to a trivial fixed point.

The vanishing of the correlation length at infinite temperature where the spins are totally uncoupled, or at zero temperature where they are frozen in a ground state is of little physical interest here and we shall ignore it. However, we have clearly made good our promise to prove

$$T_C = T^*, \quad (5.28)$$

i.e., to show that the critical temperature is located at the fixed point. We will now show that the properties of the renormalization group in the vicinity of the fixed point determine the values of the critical exponents.

### 5.3.1 Linearization about a fixed point

A renormalization transformation is, in general, a non-linear transformation as evident, for example from (5.11)-(5.13), but in the close vicinity of a fixed point we should be able to linearize it on the assumption that it behaves sufficiently smoothly. In the present context this merely means replacing the curved plot of  $\mathcal{R}(T)$  near  $T^*$  by its tangent at  $T^*$ . Writing, as before

$$t = \frac{(T-T_c)}{T_c} = \frac{(T-T^*)}{T^*}, \quad (5.29)$$

it follows that after renormalization the temperature deviation will be given by

$$t' \equiv t^{(1)} \approx \Lambda_1(b)t, \quad (5.30)$$

for small enough  $t$ , where  $\Lambda_1(b)$  is the slope of the tangent, which, as has been indicated will depend explicitly on the spatial rescaling factor  $b$ . To see this suppose one iterates twice so obtaining

$$t'' \equiv t^{(2)} \approx \Lambda_1(b)\Lambda_1(b)t. \quad (5.31)$$

Clearly this should be quite equivalent to transforming with a spatial rescaling factor  $b^2$ . Thus we conclude that one must also have

$$t^{(2)} \approx \Lambda_1(b^2)t, \quad (5.32)$$

from which we see that

$$\Lambda_1(b)\Lambda_1(b) = \Lambda_1(b^2). \quad (5.33)$$

We learn from this that  $\Lambda_1$  must depend on  $b$  in a rather special way, namely as

$$\Lambda_1(b) = b^{\lambda_1}, \quad (5.33a)$$

where  $\lambda_1$  is some constant (independent of  $b$ ). If the renormalization transformation is iterated  $\ell$  times one clearly obtains

$$t^{(\ell)} \approx \Lambda_1^\ell t = b^{\lambda_1 \ell} t, \quad (5.34)$$

while the effect on the correlation length follows from (5.26) as

$$\xi(t) \approx b^\ell \xi(b^{\lambda_1 \ell} t). \quad (5.35)$$

We have been explicit in these equations about the fact that the behavior stated really holds only asymptotically close to the critical point within the regime where the linearization represents a good approximation.

Now  $\ell$ , the number of iterations, is quite arbitrary and so we may select its value in a way which procures a major simplification. Specifically if we choose  $\ell$  to satisfy

$$b^\ell = t^{-1/\lambda_1}, \quad (5.36)$$

the flow relation for the correlation length becomes

$$\xi(t) \approx t^{-1/\lambda_1} \xi(1) = \frac{\text{const.}}{t^{1/\lambda_1}}. \quad (5.37)$$

This evidently matches the power law behavior,  $\xi \approx 1/t^\nu$ , which we expect to see near a critical point! On comparing exponents we make the identification

$$\nu = \frac{1}{\lambda_1}. \quad (5.38)$$

One can carry out a precisely similar analysis, based on (5.20), for the free energy. The result that emerges (on choosing  $C$  so that  $f$  vanishes at the fixed point) is

$$f(t) \approx b^{-d\ell} f(b^{\lambda_1 \ell} t), \quad (5.39)$$

from which, again choosing  $b$  to satisfy (5.36), one obtains

$$f(t) \approx t^{d/\lambda_1} f(1) = t^{d\nu} f(1). \quad (5.40)$$

This we may compare with the standard critical behavior for the free energy which we recall from (3.25) is  $f \approx A_0 t^{2-\alpha}$ . Hence we deduce the hyperscaling relation

$$2-\alpha = \frac{d}{\lambda_1} = d\nu, \quad (5.41)$$

first introduced heuristically in Sec. 3.8.

### 5.3.2 A second variable and scaling

If the magnetic field  $H$  is no longer constrained to vanish, the renormalization group operator,  $\mathcal{R}_b$ , acts in a more complicated non-linear fashion to generate  $T'$  and  $H'$  from  $T$  and  $H$ . We can express this fact either in terms of the pair of coupled recursion relations

$$T' = \mathcal{R}_T(T, H), \quad (5.42)$$

$$H' = \mathcal{R}_H(T, H), \quad (5.43)$$

or as the "vector" recursion relation

$$\begin{pmatrix} T \\ H \end{pmatrix} \Rightarrow \begin{pmatrix} T \\ H \end{pmatrix}' = \mathcal{R}_b \begin{pmatrix} T \\ H \end{pmatrix}. \quad (5.44)$$

Note that we may ignore the "constant" term  $C$  because its flow, while depending on  $T$  and  $H$ , cannot itself have any influence on  $T$  and  $H$ , since it merely represents an additive contribution to the Hamiltonian and, thence, to the free energy but does not affect the coupling or spin configurations in any way. [This can, of course, be seen explicitly in (5.11)-(5.13)]. On the other hand, in neglecting other possible variables and focusing just on  $T$  and  $H$  we are presenting what, in our quantum-mechanics analogy, might be termed only a "single-particle picture" rather than a many-particle theory which, in quantum mechanics, would entail discussion of Fermi and Bose statistics, and so on. To the extent that the nearest-neighbor one-dimensional Ising model can be treated correctly within this limited context it can be regarded as the "hydrogen atom" of critical phenomena; however, as in most of chemistry and physics, it will prove essential to move beyond the hydrogen atom to approach the most interesting problems!

With these provisos in mind, let us, as before, assume the existence of a nontrivial fixed point  $(T^*, H^*)$  which, from the symmetry of the magnetic

Hamiltonian, should occur at  $H^* = 0$ . By the previous arguments, this fixed point will again prove to be the critical point since we still obtain

$$\xi[\bar{H}^*] \equiv \xi^* = b\xi^* = \infty. \quad (5.45)$$

Near this fixed point we may linearize the recursion relations which yields

$$\begin{pmatrix} \Delta T \\ \Delta H \end{pmatrix}' \approx \tilde{L} \begin{pmatrix} \Delta T \\ \Delta H \end{pmatrix}, \quad (5.46)$$

where the linear operator,  $\tilde{L}$  is now the matrix

$$\tilde{L} = \begin{bmatrix} \frac{\partial \mathcal{R}_T}{\partial T} & \frac{\partial \mathcal{R}_H}{\partial T} \\ \frac{\partial \mathcal{R}_T}{\partial H} & \frac{\partial \mathcal{R}_H}{\partial H} \end{bmatrix}^*, \quad (5.47)$$

the derivatives being evaluated at the fixed point, while  $\Delta T = T - T^*$  and  $\Delta H = H - H^*$  denote the deviations of  $T$  and  $H$  from their fixed point values.

Now the  $2 \times 2$  matrix  $\tilde{L}$  will have two eigenvalues

$$\Lambda_1 = b^{\lambda_1} \quad \text{and} \quad \Lambda_2 = b^{\lambda_2}, \quad (5.48)$$

with associated eigenvectors say,  $g_1$  and  $g_2$  in terms of which we may expand the deviation vector as

$$\begin{pmatrix} \Delta T \\ \Delta H \end{pmatrix} = h_1 g_1 + h_2 g_2 = h_1 \begin{pmatrix} q_{11} \\ q_{12} \end{pmatrix} + h_2 \begin{pmatrix} q_{21} \\ q_{22} \end{pmatrix}. \quad (5.49)$$

The coefficients,  $h_1$  and  $h_2$ , can evidently stand in for  $\Delta T$  and  $\Delta H$  which, in turn, represent the deviations from criticality since, in this simplified "Bohr-Sommerfeld" treatment, the critical point is at the fixed point. The parameters  $h_1$  and  $h_2$  are therefore called the critical fields or, in a somewhat more general context, the linear scaling fields. In general we must expect to find, by solving (5.49), that  $h_1$  and  $h_2$  are linear combinations of  $\Delta T$  and  $\Delta H$ : this would, for example, be the case at the critical point of a fluid where  $H_1$  in particular, must be replaced by a particular combination of  $\Delta p = (p - p_c)$ , the pressure deviation, and  $\Delta T = (T - T_c) \propto t$ . In the case of a simple ferromagnet, such as we have in mind, symmetry under  $H \Rightarrow -H$  dictates that  $\tilde{L}$  is a diagonal matrix and hence that we have

$$h_1 = t \quad \text{and} \quad h_2 = H, \quad (5.50)$$

where we have chosen a convenient normalization for the eigenvectors.

On iterating the linearized renormalization group transformation (5.46)  $l$  times

we obtain

$$t^{(\ell)} \approx b^{-\lambda_1 \ell} t \quad \text{and} \quad H^{(\ell)} \approx b^{-\lambda_2 \ell} H, \quad (5.51)$$

provided that  $T^{(\ell)}$  and  $H^{(\ell)}$  remain in the linear region. If the iterations are repeated too often then, ultimately, the flow of  $T$  and  $H$  will become non-linear and the forms (5.51) break down. This can be seen explicitly by using the recursion relations derived for the one-dimensional Ising model for which the overall flow pattern is shown in Fig. 5.3. Notice that this actually displays a whole line of trivial fixed points! The nontrivial fixed point of interest here occurs at  $T^* = 0$  corresponding to the fact, discussed in Sec. 4.3, that critical point behavior occurs at  $T_c = H_c = 0$ .

In the linear region near the critical point the flow of the free energy is now given by

$$f(t, h) \approx b^{-d\ell} f(b^{-\lambda_1 \ell} t, b^{-\lambda_2 \ell} H), \quad (5.52)$$

(where, again,  $C$  has been chosen so that  $f$  vanishes at the fixed point). On making use of the freedom to choose the value of  $\ell$  we may set

$$b^\ell = \left( \frac{t}{t^\dagger} \right)^{-1/\lambda_1}, \quad (5.53)$$

where  $t^\dagger$  is a suitably small, fixed reference temperature, selected to keep the iterations within the linear regime. Thus we obtain the relation

$$f(t, h) \approx \left( \frac{t}{t^\dagger} \right)^{d/\lambda_1} f\left( t^\dagger, \frac{H}{(t/t^\dagger)^{\lambda_2/\lambda_1}} \right). \quad (5.54)$$

But since  $t^\dagger$  is now just a fixed parameter, this result corresponds exactly to our original scaling ansatz for the free energy, namely

$$f(t, h) \approx A_0 t^{2-\alpha} \Upsilon\left(D \frac{H}{t^\Delta}\right). \quad (5.55)$$

Comparison yields the exponent identification

$$2-\alpha = \frac{d}{\lambda_1}, \quad \text{and} \quad \Delta = \frac{\lambda_2}{\lambda_1}. \quad (5.56)$$

Thus we see how the renormalization group eigenvalues,  $\lambda_1$  and  $\lambda_2$ , at the appropriate nontrivial fixed point determine the critical exponents! At the same time scaling is implied just by the form of the transformation. The non-universal amplitudes,  $A_0$  and  $D$ , are also easily expressed in terms of  $t^\dagger$ ,  $\lambda_1$  and  $\lambda_2$ , while the scaling function itself is given formally by

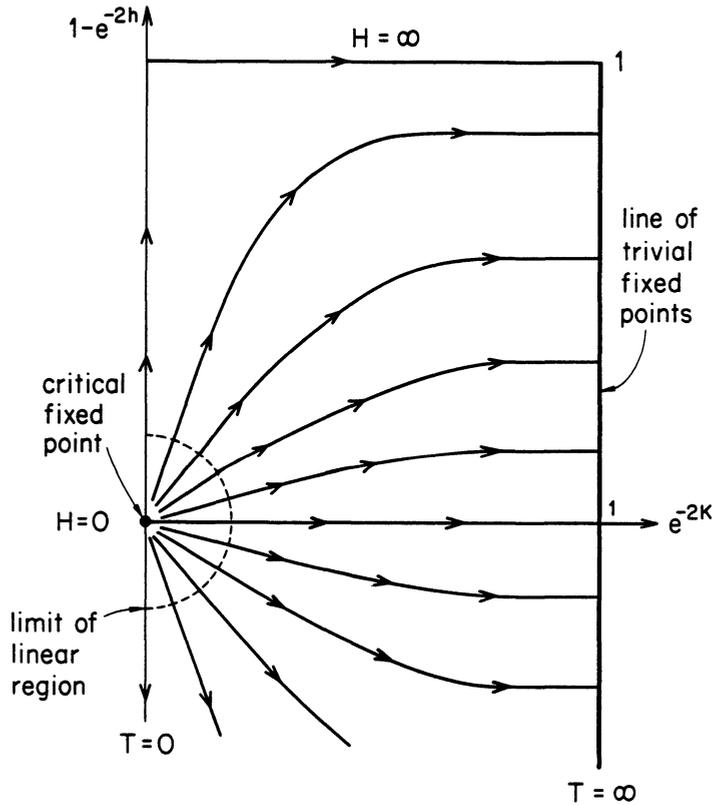


Fig. 5.3 Overall flow pattern in the  $(T, H)$  plane for the decimation or dedecoration renormalization group for the one-dimensional Ising model (based on Nelson and Fisher (1975) loc. cit.). Recall that  $K = J/k_B T$  and  $h = H/k_B T$ . The dashed curve delimits, approximately, the region over which a linearization of the renormalization group (in this case in the variables  $x = \exp(-4J/k_B T)$  and  $H$ ) is justifiable.

$$Y(y) = f(t^\dagger, y), \quad (5.57)$$

where  $t^\dagger$  must evidently be chosen sufficiently small to eliminate the effects of the neglected nonlinearities. Sometimes the scaling function may be calculated by a "matching" procedure which involves iterating  $\mathcal{TR}$  sufficiently many times that a noncritical region is reached where the renormalized free energy can be matched, to sufficient accuracy, onto results obtained from some other theory, such as mean field theory or perturbation theory which can be regarded as valid away from criticality.

At this point the reader will find it a very instructive exercise to return to the exact recursion relations (5.11)-(5.13) for the linear Ising chain and work explicitly through the chain of reasoning leading first to the zero-field fixed point and evaluation of the critical exponents  $\alpha$ , and  $\nu$  and then through the two-variable situation to the scaling behavior (5.54) and the gap exponent  $\Delta$ . It will be found that the nontrivial fixed point occurs at  $T^* = H^* = 0$ . Because this is a zero-temperature fixed point one finds it appropriate to use the variable  $x = e^{-4K}$  in place of  $T$ , in terms of which the recursion relations are readily linearized. In this way all the critical features derived in Sec. 4.3 are recaptured correctly without the need of solving exactly for the full free energy: that, of course, is what the renormalization group is all about! Details will be found in Nelson and Fisher<sup>17</sup> but we quote the exponents

$$2 - \alpha_x = \nu_x = \frac{1}{\lambda_x} = \frac{1}{2} \quad \text{and} \quad \Delta_x = \frac{\lambda_h}{\lambda_x} = \frac{1}{2}, \quad (5.58)$$

where the subscript  $x$  denotes the use of  $x$  in place of  $t$  in the exponent definitions and scaling forms, while  $\lambda_h$  is merely an alternative notation for the eigenvalue  $\lambda_2$ . The scaling function is found to be simply

$$Y(y) = (1 + y^2)^{\frac{1}{2}}, \quad (5.59)$$

the nonuniversal amplitudes then being  $A_0 = D = 1$ .

#### 5.4 General Renormalization Groups

We have used the one-dimensional Ising model to introduce some of the most important aspects of renormalization group theory at an initial level. However, we have presented no explanation of the observed universality of critical phenomena. Nor have we shown how one might construct a renormalization group transformation,  $\mathcal{TR}$ , for systems of higher dimensionality or with other types of local variables than the simplest Ising spins. Neither have we seen how to calculate explicitly for

more complex systems. Accordingly, we will now resurvey the terrain, but from a general perspective, presenting, as it were the Schrodinger picture of critical phenomena.

#### 5.4.1 The space of Hamiltonians

Let us start by listing some of the essential attributes of, and some of the important assumptions we will make (or that should be proved) about an effective renormalization group transformation. A crucially important point is the need, in formulating  $\overline{\mathcal{R}}$ , for a "large" space,  $\overline{\mathcal{H}}$ , of Hamiltonians. Historically, this aspect was rather late in being generally recognized; it was K. G. Wilson who first emphasized it strongly. A simple example serves to illustrate why and how this need arises. Consider the double Ising chain or two-layer lattice which, pictorially, constitutes a ladder. What is the effect on the corresponding Hamiltonian of a  $b=2$  decimation which eliminates alternate pairs of spins? In the absence of a

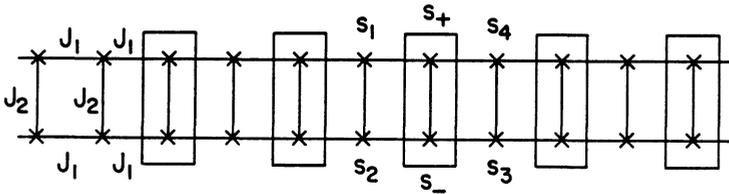


Fig. 5.4 A double-chain or two-layer Ising model ladder with nearest neighbor interactions of strength  $J_1$  along the chains and  $J_2$  between the chains. The boxed pairs of spins are summed over and thence eliminated in a  $b=2$  decimation transformation.

magnetic field the original Hamiltonian is specified, as shown in Fig. 5.4, by three interaction parameters, namely,  $J_1$ , the coupling between nearest neighbors along one chain,  $J_2$ , the cross-chain coupling between adjacent spins on opposite chains, and by  $C$ , the additional "constant" parameter that was introduced earlier in treating the simple Ising chain. Carrying out the partial trace clearly results in a new double chain, and we need to match the Hamiltonians for this renormalized system and for the original system on sets of four untransformed spins like  $s_1, s_2, s_3$  and  $s_4$  in Fig. 5.4. Now there are 16 possible configurations of these four spins and these give rise to 16 matching equations. For  $H=0$  a number of these equations turn out to be equivalent. Nevertheless, as each reader should convince him or herself, it is

quite impossible to achieve matching using only three renormalized coupling constants  $C'$ ,  $J_1'$ , and  $J_2'$  as would be expected for a simple ladder. Rather it proves essential to introduce two more new parameters, namely,  $J_3'$ , for pair couplings like  $s_1 s_3$  and  $s_2 s_4$ , and  $J_4'$ , for a quartic coupling term  $s_1 s_2 s_3 s_4$  as illustrated in Fig. 5.5. Thus we establish the need for an enlarged space,  $\mathcal{H}$ ,

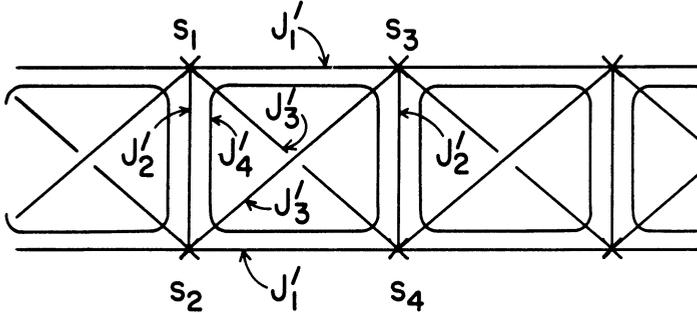


Fig. 5.5 A "braced ladder" of Ising spins with four-spin couplings resulting from renormalizing the simple Ising ladder of Fig. 5.4 by a decimation transformation.

of in this case five-parameter Hamiltonians. In the "initial", physically given Hamiltonian two of these,  $J_3$  and  $J_4$ , just "happen" to vanish identically! After renormalization, however, they necessarily appear.

In general, then, one must allow for an indefinitely large space  $\mathcal{H}$  of Hamiltonians  $\mathcal{H} \equiv (C, K_1, K_2, K_3, \dots)$ , in order to provide a reasonable chance for a useful renormalization group to exist. The Hamiltonians will be characterized by coupling parameters  $C$ ,  $K_1 \equiv J_1/k_B T$ ,  $K_2 \equiv J_2/k_B T$ , etc. which, in general, will be infinite in number. For this reason renormalization group problems tend to be difficult, and as yet, there are not many that have been solved exactly or analyzed by rigorous methods.

#### 5.4.2 Renormalization group desiderata

A renormalization group  $\mathcal{R}_b$  for a space  $\mathcal{H}$  of Hamiltonians should satisfy the following requirements:

A. Existence. There should, in the first place, clearly be a well-defined transformation, or mapping,

$$\bar{\mathcal{H}} \Rightarrow \bar{\mathcal{H}}' = \mathcal{R}_b[\bar{\mathcal{H}}], \quad (5.60)$$

which, in particular, remains unambiguous and well-defined in the thermodynamic limit  $N \rightarrow \infty$ . (See also below.)

B. Elimination. In the process of making the transformation there should be a reduction in the number (or density) of the original degrees of freedom, so that we can write, as before,

$$N \Rightarrow N' = N/b^d, \quad (5.61)$$

where  $b$  is the spatial rescaling factor.

C. Spatial locality. The transformation should not be so drastic that it mixes up the local degrees of freedom, the spins, in a hopelessly haphazard way! More concretely, one should be able to identify the same regions of space and associated local variables before and after the transformation, although, of course, spatial distances will have been changed: two regions of space originally separated by a distance  $\bar{x}$  will be brought closer together by a factor  $b$  after the transformation,

$$\bar{x} \Rightarrow \bar{x}' = \frac{\bar{x}}{b}, \quad (5.62)$$

thus preserving the overall density of degrees of freedom (which is what, basically, fixes distance scales).

Hand in hand with this rescaling of space goes the transformation of the correlation length according to

$$\xi \Rightarrow \xi' \approx \xi/b. \quad (5.63)$$

However, this particular relation must, in general, be regarded as mainly heuristic since the true transformation for  $\xi$ , especially away from criticality, must depend on the details of the particular definition of correlation length which is adopted.<sup>18</sup>

In momentum space the effect of  $\mathcal{R}_b$  is to enlarge wave-vectors by a factor  $b$ , so that

$$q \Rightarrow q' = bq. \quad (5.64)$$

D. Unitarity. Thermodynamics should be preserved by a renormalization group transformation. In other words, the two Hamiltonians should give rise to the equivalent thermodynamic functions under proper transformations of the thermodynamic

fields or the couplings. In particular the unitarity relation

$$Z_N[\bar{\mathcal{H}}] = Z_N'[\bar{\mathcal{H}}'], \quad (5.65)$$

preserves the total partition function and from this we obtain the flow equation or recursion relation for the overall free energy density, namely,

$$f[\bar{\mathcal{H}}] = b^{-d} f[\bar{\mathcal{H}}'], \quad (5.66)$$

as already demonstrated in (5.19).

E. Smoothness and uniformity. In employing a renormalization group transformation it is normally essential to assume that the transformation is smooth in the sense that if  $\mathcal{H} \Rightarrow \mathcal{H}'$  and  $\mathcal{H} + \delta\mathcal{H} \Rightarrow \mathcal{H}' + \delta\mathcal{H}'$  then as  $\delta\mathcal{H} \rightarrow 0$  one has  $\delta\mathcal{H}' \rightarrow 0$ ; or, more strongly, that  $\delta\mathcal{H}'$  becomes proportional to  $\delta\mathcal{H}$  so that a first derivative exists; and so on for one or more higher derivatives. Again, such smoothness is normally assumed to hold uniformly over interesting regions of the Hamiltonian space  $\mathcal{H}$  and to apply, in particular, to the flow equation for the free energy where, furthermore, one trusts that one is entitled to neglect the differences  $f_N[\mathcal{H}] - f_\infty[\mathcal{H}]$  and their derivatives, etc., in the thermodynamic limit  $N \rightarrow \infty$ . These properties are not obviously guaranteed and really need thought and justification as the specific cases arise. (Indeed, smoothness has been seriously questioned by Griffith's and Pearce<sup>19</sup> for certain types of renormalization group transformation.)

F. Aptness or focusability. For any given Hamiltonian or class of Hamiltonians there is not just one renormalization group - "the renormalization group" as some people say - but rather there are many that might be introduced, and one must question, for example, whether the process is best carried out in real space or momentum space and so on. A "good" renormalization group must be "apt" or appropriate for the problem at hand, and it must, in particular, "focus" properly on the critical phenomena of interest. To this end it is sometimes necessary to introduce additional devices to make the renormalization group work satisfactorily. An important instance is provided by spin rescaling in which the (continuous) spin variables undergo the transformation

$$\vec{s} \Rightarrow \vec{s}' = \vec{s}/c, \quad (5.67)$$

where  $c$ , the spin rescaling factor, may have to be chosen appropriately as some function  $c = c(b)$  of the rescaling factor  $b$  (or, even, as  $c[\bar{\mathcal{H}}; b]$ ). We will see concretely how this need arises in Section 6 when the momentum shell renormalization group is used to generate the  $\epsilon = 4-d$  expansion for critical exponents. However,

the significance of  $c$  can be seen more generally in the context of an important special class of renormalization groups which we characterize as quasi linear.

### 5.4.3 Quasi-linear renormalization groups

Spin rescaling has an intimate connection with the correlation functions as we now show. The spin-spin correlation function for two points, say  $\underline{x}$  and  $\underline{x}'$ , will depend on  $\underline{x}$  and also on the magnetic field and interaction terms in the Hamiltonian, so we may write quite generally

$$\langle s_{\underline{x}} s_{\underline{x}'} \rangle = G[\underline{x}; \overline{\mathcal{H}}]. \quad (5.68)$$

Now a renormalization group transformation not only changes  $\underline{x}$  and  $\overline{\mathcal{H}}$  but also involves some definition of the renormalized spin variables and their relation to the original spin variables. In the decimation transformation this simply amounted to a re-identification (or relabelling) of the original spins. More generally, however, the relation between  $s_{\underline{x}}$  and  $s'_{\underline{x}'}$  may be, and usually will be more complex. Consequently the transformation law for  $G[\underline{x}; \overline{\mathcal{H}}]$  is not necessarily simple. In the case of a quasi linear renormalization group, however, an identification such as (5.67) holds so that, in particular, the pair spin correlation function has the transformation law

$$G[\underline{x}; \overline{\mathcal{H}}] = c^2 G[\underline{x}'; \overline{\mathcal{H}}']. \quad (5.69)$$

The factor  $c^2$  appears simply because each spin in the definition (5.68) is to be rescaled.

Now, granting such a relation, consider the situation at a nontrivial fixed point which, by definition satisfies

$$(\overline{\mathcal{H}}^*)' = \mathcal{R}_b[\overline{\mathcal{H}}^*] = \overline{\mathcal{H}}^*, \quad (5.70)$$

so that (5.69) yields, for the fixed point correlations

$$G^*(\underline{x}) = (c^*)^2 G^*(\underline{x}/b). \quad (5.71)$$

Since  $b$  is essentially arbitrary this represents a functional equation for  $G^*$  which has the unique solution

$$G^*(\underline{x}) \approx \frac{D}{x}, \text{ with } c^* = c[\overline{\mathcal{H}}^*] = b^{-\omega}, \quad (5.72)$$

where  $D$  and  $\omega$  are constants (independent of  $b$ ). However, such power law behavior is just what is to be expected at a fixed point which represents the critical point of a system. Thus we can make the identification

$$\omega = \frac{1}{2} (d-2+\eta), \quad (5.73)$$

where  $\eta$  is the critical point decay exponent introduced in (3.48).

This conclusion can be restated in another way: in order to obtain a nontrivial fixed point of appropriate critical character, it is necessary to adjust  $c$  (at least close to  $\bar{c} = \bar{c}^*$ ) to satisfy (5.73), where  $\eta$  need not be known a priori. This is somewhat analogous to the adjustment of the energy in a Schrodinger equation for a stationary state so that the wave function satisfies proper boundary conditions, and the energy then yields the desired eigenvalue.

If one now returns to our exact decimation solution of the one-dimensional Ising model one sees that, without raising the question, we implicitly took a spin rescaling factor  $c = 1$  or  $\omega = 0$  in (5.72). Further, as mentioned, the quasilinear criterion, (5.69), was indeed satisfied. The fact that we then obtained a sensible fixed point for one-dimensional Ising criticality at  $T = 0$  was thus really a result of the "accident" that  $\eta = 1$  describes the scaling behavior of the Ising correlations via (3.52) as  $T \rightarrow 0$ : this, happily, agrees with (5.73) since  $2\omega-d+2 = 0-1+2 = 1!$  It is evident, however, that if one suspects that  $\eta$  is not simply equal to  $(2-d)$  it is inappropriate to use a quasilinear renormalization group unless one allows for a spin rescaling factor.

## 5.5 Flows, universality and scaling

The assumption of smoothness means that systems represented by Hamiltonians corresponding to nearby points in our multidimensional parameter space,  $\mathcal{H}$ , flow under renormalization to other points which also lie relatively close together. Let us apply this observation to the set of Hamiltonians representing a single physical system, say for concreteness the ferromagnet pure nickel, in the vicinity of its critical point. We will enquire into the flow trajectories generated by iteration of the renormalization group and see how this leads naturally to a concept of universality.

### 5.5.1 Universality, relevance and irrelevance

Consider Fig. 5.6 which presents a visualization of the space,  $\mathcal{H}$ , of Hamiltonians and, in particular, exhibits a "physical manifold" described by the

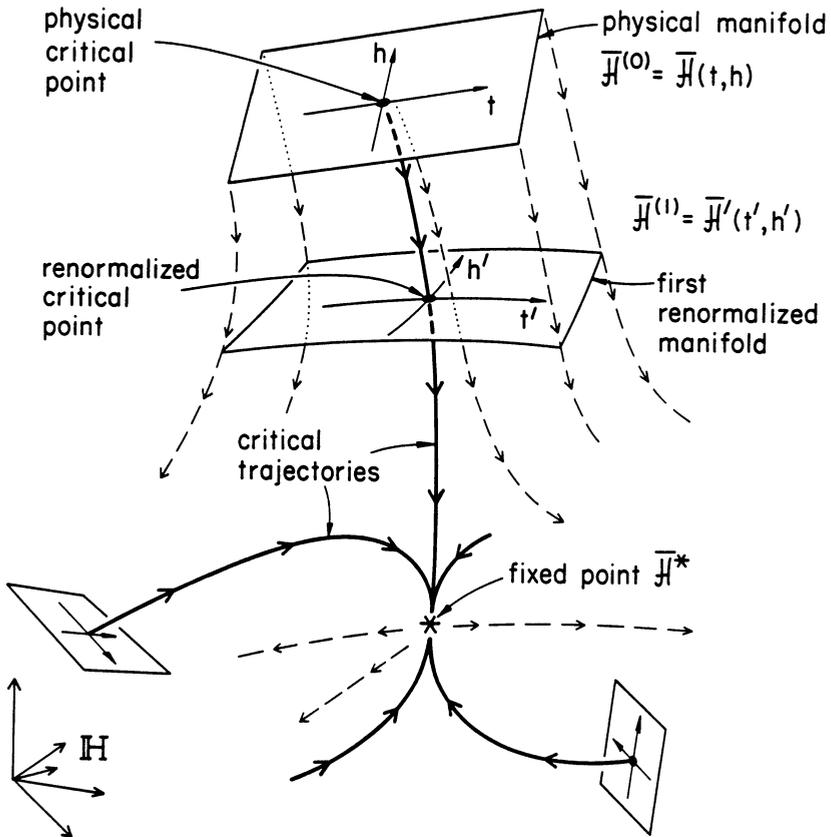


Fig. 5.6 A representation of the space of Hamiltonians,  $\bar{H}$ , showing initial or physical manifolds and subsequent renormalization group flows. Critical trajectories are shown bold: they terminate on the fixed point  $\bar{H}^*$ .

initial, unrenormalized Hamiltonians,  $\mathcal{H}_c^{(0)}(t,h)$ , corresponding, as we agreed, to nickel near its ferromagnetic critical point. At the critical point itself, when the physical fields  $t$  and  $h$  vanish, we have  $\xi = \infty$  because of the characteristic slow decay of the correlations. However, the critical Hamiltonian  $\mathcal{H}_c^{(\infty)} = \bar{\mathcal{H}}_c^{(0)}(0,0)$  is not, in general, a fixed point! After one operation of the renormalization group transformation we obtain a new manifold, representing the first-renormalized Hamiltonians  $\bar{\mathcal{H}}_c'(t',h')$ , in which is embedded the renormalized critical Hamiltonian,  $\bar{\mathcal{H}}_c' = \mathcal{R}_b[\bar{\mathcal{H}}_c]$ . It is crucial to realize that this also will be a critical Hamiltonian [and, hence, equal to  $\bar{\mathcal{H}}_c'(t'=0, h'=0)$ ]: the reason for this is simply that the flow equation (5.63) for the correlation length tells us that  $\xi_c \Rightarrow \xi_c' = \xi_c/b = \infty/b = \infty$ ! Thus under successive renormalization a line or trajectory of critical points is generated. In principle this critical trajectory might eventually fly off to infinity or it might wander around in  $\mathbb{H}$  forever, even perhaps in some sort of turbulent or chaotic motion! Nevertheless in the light of the previous examples (and further calculations to be performed) it is also very plausible to suppose that the critical trajectory eventually terminates at some fixed point  $\bar{\mathcal{H}}_c^*$  at which, of course, further renormalization produces no further motion. The critical trajectory starting from the critical point of nickel and proceeding through a sequence of critical points of renormalized forms of nickel, lies on the stable critical manifold of the fixed point  $\bar{\mathcal{H}}_c^*$  i.e., the set of all points in  $\mathbb{H}$  which are ultimately carried by the renormalization group flows into  $\bar{\mathcal{H}}_c^*$ . Evidently all points on this stable critical manifold, including the fixed point itself, correspond to systems at criticality.

Now one might start in quite a different region of parameter space corresponding, say, to iron or gadolinium, as suggested by the other initial, physical manifold indicated in Fig. 5.6. Then, perhaps, under renormalization the critical point Hamiltonians for iron and gadolinium flow to the same fixed point as before! If it happens this way, then Ni, Fe and Gd must all lie on the same critical manifold. The universality of their critical behavior then follows from this fact! To demonstrate this point consider what happens to the free energy under  $l$  successive renormalization group iterations: by (5.66) we have

$$f[\bar{\mathcal{H}}] = b^{-dl} f[\bar{\mathcal{H}}^{(l)}]. \quad (5.74)$$

We see that the behavior of  $f(t,h)$  for any  $\bar{\mathcal{H}} \equiv (t,h)$  which lies near a critical Hamiltonian is determined by the behavior of  $f[\bar{\mathcal{H}}]$  for a multiply renormalized Hamiltonian which will lie close to the fixed point. Thus the critical behavior for Ni, Fe, and Gd, and for any other systems whose critical points lie on the same manifold, will be essentially identical. In particular, because of the smoothness of the mapping all will display the same critical exponents and, furthermore, all will be described by the same scaling functions. It is only as regards the various

non-universal amplitudes that the various systems will differ.

Not all systems, of course, are expected to have critical Hamiltonians which flow to the same fixed point. For instance, suppose iron is placed under a uniaxial stress. The initial parameters, and hence the initial physical manifold, are now slightly altered. If the critical point Hamiltonian were still to flow to the same fixed point as before, we would call this uniaxial stress an irrelevant perturbation since it does not change any of the essential asymptotic critical properties. On physical grounds, however, in this particular case we suspect strongly that the critical behavior will change since the uniaxial stress should enhance parallel spin fluctuations but tend to suppress transverse fluctuations. Thus the new flow should carry  $\bar{\mathcal{H}}_c$  to a different fixed point which may be described in terms of a single-component, scalar or  $n = 1$  Ising-like order parameter, whereas the original fixed point for unstressed iron is expected to correspond to an isotropic Heisenberg-like or ( $n = 3$ )-component order parameter. Now there will be another manifold of Hamiltonians that all flow to the new Ising-like fixed point; the critical properties of systems described by these Hamiltonians will be different from those of the former set. In such a case we say the uniaxial stress represents a relevant perturbation since it causes the critical Hamiltonian to flow to a distinct, new fixed point.

The flow picture brings out clearly the idea of various universality classes. Systems which belong to the same universality class have critical Hamiltonians which flow into the same (or equivalent) fixed points. The corresponding critical manifolds can be regarded as the catchment areas or basins of attraction of the different fixed points.

### 5.5.2 Continuous flows

As we have seen, one utilizes a renormalization group transformation by iterating it, obtaining successive renormalized Hamiltonians. Accordingly it usually proves convenient to introduce a discrete flow variable  $\ell$ , which counts the iterations. It can clearly be thought of as a time-like renormalization or rescaling variable which parameterizes the flow trajectories. To this end we rewrite the spatial rescaling factor as  $b = e^\ell$ , and recall that the renormalization transformation is parametrized by  $b$  as  $\bar{\mathcal{H}}' = \mathcal{R}_b[\bar{\mathcal{H}}]$ . Quite often, however, there arise situations in which  $\ell$  can be regarded as a continuous, truly time-like flow variable. When this is so, the renormalization group equations can be written more directly as differential flow equations. Thus for the Hamiltonian itself the transformation is represented by

$$\frac{d\bar{\mathcal{H}}}{d\ell} = \mathcal{G}[\bar{\mathcal{H}}], \quad (5.75)$$

where the  $\mathcal{G}$  is the infinitesimal generator for  $\mathcal{R}$ . It can thus be expressed via a limit operation as

$$\mathcal{G} = \lim_{b \rightarrow 1^+} \left( \frac{\mathcal{R}_b - 1}{b - 1} \right). \quad (5.76)$$

Since  $\bar{\mathcal{H}}$  merely represents the point  $(C, K_1, K_2, \dots)$ , the flow can also be written as a set of simultaneous differential equations of the form

$$\begin{aligned} \frac{dC}{d\ell} &= \mathcal{G}_0(C, K_1, K_2, \dots), \\ \frac{dK_1}{d\ell} &= \mathcal{G}_1(C, K_1, K_2, \dots), \end{aligned} \quad (5.77)$$

for the parameters  $C, K_1, K_2$ , etc. (If, as before,  $C$  represents the "constant" term in  $\bar{\mathcal{H}}$ , it will not actually enter in the  $\mathcal{G}_i$  for  $i > 0$ .)

### 5.5.3 The fixed point spectrum

In order to use a renormalization group to describe critical phenomena we must assume that there is an appropriate fixed point  $\bar{\mathcal{H}}^*$ . This assumption is backed up in many cases by various more-or-less detailed calculations. A few, like those for the one-dimensional Ising model, are exact but most are at best systematic approximations. However, if we follow the assumption through, powerful general conclusions follow: conversely if no proper fixed point exists we may expect scaling and other consequences to fail.

The first step, as we have seen, is linearization. To implement the procedure we take  $\bar{\mathcal{H}}$  close to the fixed point and write

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}^* + g\mathcal{Q}, \quad (5.78)$$

where  $g$  is small and  $\mathcal{Q}$  is some "operator", i.e., a partial Hamiltonian. On operating with  $\mathcal{R}_b$  and invoking the smoothness assumption, we obtain

$$\bar{\mathcal{H}}' = \mathcal{R}_b[\bar{\mathcal{H}}] = \bar{\mathcal{H}}^* + g\mathcal{L}_b\mathcal{Q} + O(g^2), \quad (5.79)$$

where  $\mathcal{L}_b = (\delta\bar{\mathcal{H}}'/\delta\bar{\mathcal{H}})$  is a linearized renormalization group operator. As a linear operator, it can be expected to have a spectrum of eigenvalues  $\Lambda_j(b)$  and associated "eigenvectors",  $\mathcal{Q}_j$ , which are "operators" or partial Hamiltonians. Sometimes the  $\mathcal{Q}_j$  are called critical densities or scaling operators, etc. They are determined by the eigenvalue equation

$$\mathcal{L}_b \mathcal{Q}_j = \Lambda_j(b) \mathcal{Q}_j. \quad (5.80)$$

Each of the eigenvalues should be expressible in the form

$$\Lambda_j(b) = b^{\lambda_j}, \quad (5.81)$$

where the individual  $\lambda_j$ 's are independent of  $b$ . (This reflects the semigroup property,  $\mathcal{R}_{b_1 b_2} = \mathcal{R}_{b_1} \mathcal{R}_{b_2}$ , of the renormalization group transformations.) Typically, one can make the identifications

$$\mathcal{Q}_1 = \mathcal{E}, \quad \mathcal{Q}_2 \equiv \Psi, \quad (5.82)$$

where  $\Psi$  denotes the order parameter and  $\mathcal{E}$  the energy (see also below).

If we assume the eigenvectors form a complete set, or at least a sufficiently complete basis in some asymptotic sense, we may expand  $\bar{\mathcal{K}}$  in terms of them as

$$\bar{\mathcal{K}} = \bar{\mathcal{K}}^* + \sum_j g_j \mathcal{Q}_j + \dots \quad (5.83)$$

Acting on  $\bar{\mathcal{K}}$  with  $\mathcal{R}_b$  then yields

$$\bar{\mathcal{K}}' = \bar{\mathcal{K}}^* + \sum_j g_j \Lambda_j \mathcal{Q}_j + O(g_j^2, g_1 g_j), \quad (5.84)$$

and on iterating  $\ell$  times we find

$$\bar{\mathcal{K}}^{(\ell)} = \bar{\mathcal{K}}^* + \sum_j g_j \Lambda_j^\ell \mathcal{Q}_j + O(g_j^2, g_1 g_j). \quad (5.85)$$

The  $g_j$  are called critical fields or linear scaling fields. Evidently we may express the  $\ell$ -renormalized field as

$$g_j^{(\ell)} \approx g_j \Lambda_j^\ell = b^{\ell \lambda_j} g_j, \quad (5.86)$$

where we must write  $\approx$  ("asymptotically equals") in place of  $=$  because we are neglecting the higher order terms in (5.84) and (5.85). Now as  $\ell$  increases there are three possible courses for  $g_j^{(\ell)}$ :

(a) If  $\lambda_j > 0$  then we have  $\Lambda_j > 1$  and  $g_j^{(\ell)}$  grows rapidly larger, carrying the system away from the fixed point and, hence, away from the corresponding criticality. In accord with our previous discussion, such  $g_j$  are called relevant fields and the associated  $\mathcal{Q}_j$  are called relevant operators. At a normal critical point we know that criticality is destroyed by varying the temperature, which couples to the energy,  $\mathcal{E}$ , from  $T_c$  or by changing the ordering field,  $h$ , which couples to  $\Psi$ , from its critical value,  $h = 0$ . Thus we expect to find two relevant

scaling fields which, reflecting (5.82), may be identified as  $g_1 \equiv t$  and  $g_2 \equiv h$ .

(b) If  $\lambda_k < 0$  one has  $\Lambda_k < 1$  and  $g_k^{(\ell)}$  shrinks steadily to zero. Ultimately, therefore, it should be possible to ignore such fields. For this reason, again in concordance with the earlier discussion,  $g_k$  is called an irrelevant variable or field in such a case and the associated  $\mathcal{Q}_k$  is an irrelevant operator. If the relevant fields are all set to zero the flows will take  $\mathcal{K}$  to the fixed point, so it must then lie on the critical manifold. Thus another way of stating universality is to note that the fixed point is insensitive to the irrelevant variables, so that systems differing from one another only with respect to irrelevant variables belong to the same universality class and are "governed" or "controlled" by the same fixed point.

(c) Finally there is the borderline where  $\lambda_m = 0$  so that  $\Lambda_m = 1$ . The corresponding  $g_m$  are called marginal variables, and neither grow nor shrink very rapidly. Rather the flow of a marginal variable must be described by

$$\frac{dg_m}{d\ell} = 0 + O(g_1 g_j, g_j^2), \quad (5.87)$$

and so is determined by terms quadratic in the fields. Thus a marginal variable varies only relatively slowly with  $\ell$ . On following through an analysis in which marginal variables feature (see e.g. in Sec. 6) one finds there are various special things that can happen, which violate the simplest scaling precept. One of the typical effects is the appearance of logarithmic correction factors, such as  $(\ln|t|)^v$ , multiplying the usual critical power laws. The ability to identify and predict such departures from straight forward scaling represents one of the powers of the renormalization group approach.

#### 5.5.4 Scaling of the free energy and hyperscaling

Let us now express the free energy in terms of the set of scaling fields  $g_1, g_2, g_3$ , etc. The flow equation for the free energy then takes on the simple asymptotic form

$$f(t, h, g_3, \dots) \approx b^{-d\ell} f(b^{\ell\lambda_1} t, b^{\ell\lambda_2} h, \dots, b^{\ell\lambda_j} g_j, \dots), \quad (5.88)$$

where we have made use of the identifications  $g_1 \equiv t$  and  $g_2 \equiv h$ .

Now we can make the previous choice for  $b$  or  $\ell$  by setting

$$b^{\ell\lambda_1} = 1/t, \quad (5.89)$$

which yields the general scaling result

$$f(t, h, \dots, g_j, \dots) \approx t^{2-\alpha} f(1, \frac{h}{t^\Delta}, \dots, \frac{g_j}{t^{\phi_j}}, \dots), \quad (5.90)$$

where the standard thermodynamic exponents are given, as before, by

$$2-\alpha = d/\lambda_1 \quad \text{and} \quad \Delta = \lambda_2/\lambda_1, \quad (5.91)$$

while the "crossover exponent" for the scaling field  $g_j$  is given by

$$\phi_j = \lambda_j/\lambda_1. \quad (5.92)$$

Now if  $\phi_j > 0$  for some  $j$ , the scaled combination  $g_j/t^{\phi_j}$  becomes large as  $t \rightarrow 0$  and so it clearly cannot be ignored: in other words  $g_j$  is another relevant variable and its presence will normally lead to crossover to different critical behavior (or, perhaps, to noncriticality as for  $t$  and  $h$ ). On the other hand, when  $\phi_k$  is negative one has

$$g_k/t^{\phi_k} \equiv g_k t^{\theta_k} \rightarrow 0 \text{ as } t \rightarrow 0, \quad (5.93)$$

and so  $g_k$  becomes inconsequential: it is an irrelevant variable. By expanding (5.90) in terms of the scaled combination  $g_k/t^{\phi_k}$ , if this is allowed, we see that such irrelevant variables can contribute "corrections-to-scaling" i.e., correction factors to leading power laws of the form  $[1 + c_j t^{\theta_j} + \dots]$ . At some slight risk therefore (in case the  $g_k$  enter in a "dangerous" way), one can thus discard all the irrelevant variables and worry only about the relevant ones. Finally, this justifies the postulate of asymptotic scaling near a critical point in terms of only a few important variables. For a standard critical point with only two relevant variables we thus recapture the scaling form

$$f(t, h, g_3, g_4, \dots) \approx t^{2-\alpha} Y\left(\frac{h}{t^\Delta}\right), \quad (5.94)$$

in considerable generality.

It is worth mentioning that in addition to the singular corrections to this asymptotic scaling form which arise from the irrelevant variables and their exponents as factors  $(1+ct^\theta+\dots)$ , one must always expect farther analytic corrections to scaling which will appear as  $t$  and  $h$  depart increasingly from criticality. At the most trivial level the "harmless" change from  $t = (T-T_c)/T_c$  to  $t' = 1-(T_c/T)$ , which is often useful theoretically and experimentally, introduces such correction terms since one has  $t' = t - t^2 + t^3 + \dots$ . More generally, on solving the recursion relations near a fixed point beyond linear order one finds that the scaling fields  $t, h, \dots, g_j, \dots$  in the scaling relation (5.90) should, for greater accuracy, be replaced by non-linear scaling fields,  $\tilde{t}, \tilde{h}, \dots, \tilde{g}_j, \dots$  which, in quadratic and higher order can couple the original fields together so that, for example, one

has  $\tilde{t} = t + a_1 t^2 + a_2 h^2 + a_3 t h^2 + \dots$ , where symmetry may dictate that certain terms are absent although the coefficients  $a_i$  here are nonuniversal. Clearly further analytic corrections to asymptotic scaling arise from this source and can be significant in practice.

Finally, let us appeal to the locality assumption for the renormalization group and recall (5.63) to obtain, for  $h = 0$ ,

$$\xi(t, 0, \dots) \approx b \xi(b^{-1} t, 0, \dots) \approx t^{-1/\lambda_1} \xi(1, 0, 0, \dots), \quad (5.95)$$

where we have used (5.89) and allowed the irrelevant variables to go to zero. Comparing this with the definition  $\xi(T) \sim t^{-\nu}$  as  $t \rightarrow 0$  ( $h=0$ ) yields again (see (5.38)) the identification

$$\nu = 1/\lambda_1, \quad (5.96)$$

which may, with good reason, be regarded as the most fundamental of the renormalization group exponent relations. However, the provisos explained after (5.63) and in Sec. 5.4.3 must be borne in mind and, more properly, one should work with the correlation flow equation (5.69) or its analogue for nonquasilinear renormalization groups.

If we combine (5.96) with (5.91) we immediately obtain the hyperscaling relation,  $d\nu = 2 - \alpha$ , first introduced in Sec. 3.8 on heuristic grounds [see (3.57) to (3.59)]. From this and the previous  $d$ -independent scaling relations, follow all the other hyperscaling relations such as (3.60), which relates  $\eta$  and  $\delta$ , (3.65) and (3.72). It is clear at this stage that hyperscaling is "built into" renormalization group theory in a rather intimate and deep way. Nevertheless hyperscaling fails, as mentioned previously, for classical or mean field theory (unless one has  $d = 4$ ); but we have already seen evidence, most concretely through the exact results for spherical models, that the classical exponent values are valid for  $d > 4$ ; furthermore, this is confirmed generally by the explicit renormalization group  $\epsilon$ -expansion analysis presented below in Sec. 6! Thus we are faced with the paradox that hyperscaling seems to be predicted very generally by renormalization group analysis but, nonetheless, fails strongly for  $d > 4$ : this issue is discussed further in Appendix D where it is resolved in a consistent way in terms of the properties of dangerous irrelevant variables.

This is also an appropriate place to caution the reader that one can encounter, in the critical spectrum of operators, certain so-called redundant operators: these appear formally in the specification of the Hamiltonian  $\overline{\mathcal{H}}$  and its flow under renormalization but the associated scaling fields have no effect on the free energy or other observable properties! As discussed by Wegner,<sup>20</sup> redundant operators may be envisaged as describing, in a continuous spin system, for example, a mere change

in the origin or scale of the spin variables (which, since all spins are eventually integrated out, cannot have an effect if all couplings, fields, etc. are changed in a covariant way as specified by the corresponding scaling field). In well-controlled practical calculations redundant operators do not normally cause problems and the reader will find little reference to them in the literature!

## 5.6 The construction of renormalization groups

The actual process of explicitly constructing a useful renormalization group is not trivial. We will only consider briefly a few particular renormalization groups, and then delve a little more deeply into one of them. A renormalization group typically involves going over from one set of local variables or spins to another set,  $\{s\}_N \Rightarrow \{s\}_{N'}$ . A rather general form for  $\mathcal{R}$  can be expressed via

$e^{\bar{\mathcal{H}}\{s\}} \Rightarrow e^{\bar{\mathcal{H}}'\{s'\}}$ , where the renormalized Boltzmann factor, which is what really matters, is defined by

$$e^{\bar{\mathcal{H}}'(s')} = \text{Tr}_N^s \{ \mathcal{R}_{N',N}(s',s) e^{\bar{\mathcal{H}}(s)} \}, \quad (5.97)$$

in which the kernel  $\mathcal{R}_{N',N}(s',s)$  has  $N$  original or unrenormalized variables,  $s$ , but a smaller number  $N'$  of renormalized variables  $s'$ . [The rescaling factor  $b$  is defined as usual via (5.61).] Now in order to meet the unitarity requirement (5.65), this kernel must satisfy the condition

$$\text{Tr}_{N'}^{s'} \{ \mathcal{R}_{N',N}(s',s) \} = 1 \text{ for all } s. \quad (5.98)$$

Two of the simplest and more fashionable renormalization groups can then be specified as follows:-

### 5.6.1 Kadanoff's block spin renormalization group

Kadanoff<sup>21</sup> was the first person to expose the intimate connection between the idea of a rescaled "block" or "cell" spin and the scaling properties of a critical point, thereby prefiguring Wilson's development of the general renormalization group approach. He was also the first to bring this particular approach to the point of being a practical computational scheme. (It might be mentioned, however, that the idea of block spins as a way of approaching critical phenomena had been proposed independently a year or so earlier by M. J. Buckingham<sup>22</sup> at one of the first conferences to explicitly recognize the unity and universality of diverse critical phenomena.) The simplest way of picturing Kadanoff's construction is to consider a

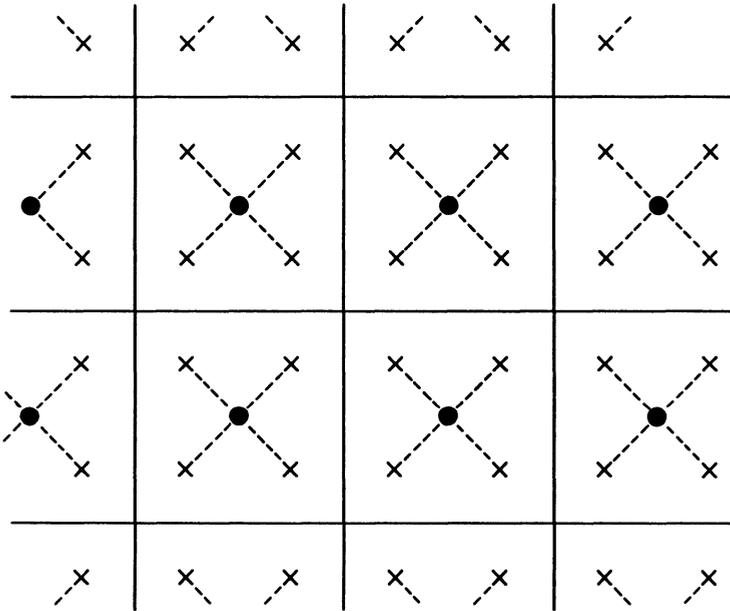


Fig. 5.7 An illustration of Kadanoff's block spin renormalization scheme for a square lattice Ising model. The original spins are denoted by crosses; the renormalized block spins are shown by solid circles. The spatial rescaling factor here is  $b = 2$ .

two-dimensional Ising model on a square lattice as illustrated in Fig. 5.7. The lattice is divided into blocks or cells each containing  $2 \times 2 = 4$  spins. On renormalization each cell of spins is replaced by a single block or renormalized spin. Thus for this renormalization group we have  $b = 2$ . There are several algebraic ways in which this replacement can be effected.

One might, for example, imagine a pair coupling of strength  $J_0$  between the block spins and the original spins in addition to the given couplings, say  $J_1, J_2$ , etc., between the original spins. A full trace is then taken over the original spins leaving a square lattice of block spins.

A crucial feature that appears directly one sums over the original spins in such a "real space" renormalization group for a lattice of dimensionality  $d > 1$  is that couplings now appear not only between first neighbor block spins but also between second neighbors, third neighbors, fourth neighbors, and so on. Worse in fact, since one is actually forced at the first step to go over to a space of Hamiltonians in which there are an infinite number of coupling constants not only between all pairs of spins but also between all triplets, all quartets, etc. Despite this inescapable complexity one can write down a formally exact expression for  $\mathcal{R}$ . In the simplest general case this is expressed by the kernel

$$\mathcal{R}_{N',N} = \prod_{\substack{\text{cells} \\ \underline{x}'}}^{N'} \frac{1}{2} [1 + cs'_{\underline{x}'} \sum_{\substack{\text{in} \\ \text{cell } \underline{x}'}}^{b^d} s_{\underline{x}'}], \quad (5.99)$$

where  $c$  serves as a spin rescaling factor. The product runs over all the blocks or cells; the sum runs over all the  $b^d$  spins in a given cell. One can also check that this yields a quasilinear transformation (as discussed in Sec. 5.4.3).

Although (5.99) is a neat closed formula, it certainly does not mean that the problem is solved! In fact the best that can be done (unless  $d=1$ ) in order to actually implement this renormalization group, i.e., to relate the new renormalized couplings explicitly to the old ones, is to invoke some approximation scheme. Unfortunately, the methods of approximation normally used entail truncating the number of interactions at each stage of renormalization to some finite number of more-or-less short range coupling terms. This is a fairly uncontrollable method of approximation, and really useful new results from this renormalization scheme have not been very plentiful.

### 5.6.2 Niemeijer and van Leeuwen's majority rule

This method is most frequently applied to plane triangular lattices but can be adapted to other geometries and dimensionalities. In the simplest case triangles of

adjacent spins,  $s$ , are formed into blocks and associated with a block spin,  $s'$ . Now in any block of three Ising spins,  $s_1$ ,  $s_2$ , and  $s_3$  say, at least two will always be pointing in the same direction. The transformation rule then states that the corresponding block spin,  $s'$ , points in the direction of the majority! The corresponding kernel can be written, with rescaling factor given by  $b^d = 3$ , as

$$\mathcal{R}_{N',N} = \prod_{\substack{\text{cells} \\ \underline{x}'}} \frac{1}{2} [1 + c_0 s'_{\underline{x}'} \sum_{\underline{x}} s_{\underline{x}} + c_1 s'_{\underline{x}'} \prod_{\underline{x}} s_{\underline{x}}], \quad (5.100)$$

where the sum and second product run over the three spins in the cell  $\underline{x}'$ ; the coefficients are fixed by  $c_0 = -c_1 = \frac{1}{2}$ . It is easily seen that this transformation is not quasilinear so that it is not necessary to adjust  $c_0$  or  $c_1$  to achieve a condition such as (5.72) and (5.73). One may be tempted, however, to try different values for  $c_0$  and  $c_1$  and to try to optimize their choice in some way. However, the exact transformation again necessarily involves an infinite number of coupling constants so some sort of approximation scheme must be used. Many possibilities arise and it is hard to find a truly systematic procedure since no small parameter presents itself. Variational criteria for choosing the optimal  $c_0$  and  $c_1$  have been explored but they cannot be relied on to yield correct final results.

### 5.6.3 Wilson's momentum shell integration

This renormalization group, which is particularly important since it turns out to allow a systematic expansion procedure, is designed for or, perhaps more fairly, requires continuous spins,  $\vec{s}_{\underline{x}} \equiv (s_{\underline{x}}^\mu)$  with  $-\infty < s_{\underline{x}}^\mu < \infty$  as discussed in Sec. 4.7. If, for simplicity, one considers a  $d$ -dimensional hypercubic lattice of spacing  $a$  (i.e., a square lattice for  $d = 2$ , simple cubic for  $d = 3$ , etc.) one can introduce the associated Fourier transformed spin variables

$$\hat{s}_{\underline{q}} = \sum_{\underline{x}} e^{i\underline{q} \cdot \underline{x}} \vec{s}_{\underline{x}}, \quad (5.101)$$

where the wave vector  $\underline{q}$  runs over the appropriate first Brillouin zone of the reciprocal space lattice: this can be expressed by

$$|q_x|, |q_y|, \dots \leq q_\Lambda = \pi/a, \quad (5.102)$$

where  $q_\Lambda$  represents a momentum space cutoff which, of course, simply reflects the underlying lattice structure. The situation for a square lattice is illustrated in Fig. 5.8. For the original spins in real space one has, reciprocally,

$$\vec{s}_{\underline{x}} = \frac{1}{N} \sum_{\underline{q}} e^{-i\underline{q} \cdot \underline{x}} \hat{s}_{\underline{q}}. \quad (5.103)$$

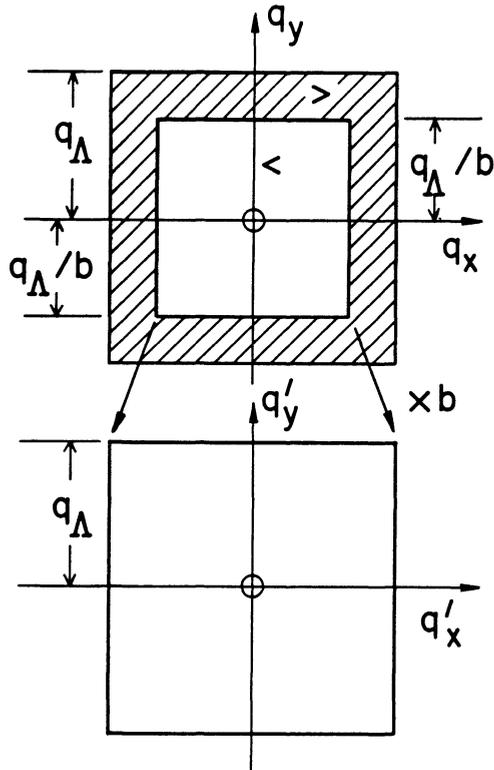


Fig. 5.8 Momentum space for a square lattice illustrating the construction of an inner zone, marked  $<$ , and an outer zone or shell marked  $>$  (upper part of figure). After integrating over spin variables with momenta in the shell, the inner zone is expanded by a factor  $b$  to form the new, renormalized Brillouin zone.

In this way any reduced Hamiltonian  $\overline{\mathcal{H}}(\hat{s}_{\underline{x}}^{\rightarrow})$ , expressed in terms of the real space (or lattice) spins, can be re-expressed precisely in terms of the Fourier spins,  $\hat{s}_{\underline{q}}$  as  $\overline{\mathcal{H}}(\hat{s}_{\underline{q}})$ . Likewise the trace operation

$$\text{Tr}_N^{\underline{x}} = \prod_{\underline{x}} \prod_{\mu=1}^n \int_{-\infty}^{\infty} ds_{\underline{x}}^{\mu}, \quad (5.104)$$

becomes simply

$$\text{Tr}_N^{\underline{g}} = \prod_{\underline{g}} \prod_{\mu=1}^n \int_{-\infty}^{\infty} d\hat{s}_{\underline{g}}^{\mu}. \quad (5.105)$$

Now, motivated by the idea that it is the low momentum or long wavelength fluctuations that are of most importance for critical phenomena, while the short wavelength, high momentum fluctuations are less crucial, Wilson divides the Brillouin zone into two regions as shown in Fig. 5.8. In the inner region, which we will indicate by a superscript  $\langle$ , the wavevectors  $\underline{q}$  satisfy

$$|q_x|, |q_y|, \dots < q_{\Lambda}/b, \quad (5.106)$$

while the remaining, outer region constitutes a "momentum shell" of thickness  $\Delta q = (1-b^{-1})\pi/a$  which can, if convenient, be chosen infinitesimal. Now the original Hamiltonian is a function of spin variables  $\hat{s}_{\underline{q}}$  with wavevectors distributed uniformly throughout the whole zone. Let us partition these into a set,  $\{\hat{s}_{\underline{q}}^{\langle}\}$ , of all those  $N' = N/b^d$  spins with wavevectors in the inner zone, and into the remaining set,  $\{\hat{s}_{\underline{q}}^{\rangle}\}$ , of the  $(N-N')$  spins in the shell. We can then write the reduced Hamiltonians as

$$\overline{\mathcal{H}}(\hat{s}_{\underline{q}}) = \overline{\mathcal{H}}(\hat{s}_{\underline{q}}^{\langle}, \hat{s}_{\underline{q}}^{\rangle}). \quad (5.107)$$

Evaluation of the partition function requires an integration over all these spins as implied by (5.105). Instead of doing this in one step, Wilson proposes that the integration be performed in stages, starting with an integration over only the spin variables  $\{\hat{s}_{\underline{q}}^{\rangle}\}$  in the outer zone or shell. This procedure clearly embodies the physical idea that the high momentum variables play a smaller role in the critical behavior and hence may reasonably be eliminated first. (It should be stressed, however, that it is a serious over-simplification to assert that all the critical behavior occurs only at low momentum: this is not the case and is not assumed in the renormalization group approach which, on the contrary, allows properly for all contributions.)

This renormalization procedure yields a new Hamiltonian  $\overline{\mathcal{H}}'$  given by

$$e^{-\hat{\mathcal{H}}'(s'_q)} = \text{Tr}_{N, N'} \left\{ e^{-\hat{\mathcal{H}}(s_q^<, s_q^>)} \right\}. \quad (5.108)$$

In this expression we have also allowed for spin and spatial rescaling. The latter proceeds simply in accord with (5.63) and (5.64). As illustrated in Fig. 5.8, the rescaling

$$q \Rightarrow q' = bq, \quad (5.109)$$

of the wavevectors corresponds to an expansion of the inner region of the original Brillouin zone to fill out the new, renormalized zone back to the size of the original zone.

A spin rescaling is needed since it is not hard to see, by examining the transformation of the Fourier space spin correlation functions  $\hat{G}(q) = \langle \hat{s}_q \hat{s}_{-q} \rangle$ , that the renormalization group defined by (5.108) is quasilinear. Accordingly, the renormalized spins are defined via

$$\hat{s}_q \Rightarrow \hat{s}'_q = \hat{s}_q / c, \quad (5.110)$$

where, in comparing with (5.67), we have the relation

$$\hat{c} = b^d c. \quad (5.111)$$

It follows by the previous arguments that at a fixed point the Fourier spin rescaling factor is related to the exponent  $\eta$  via

$$\hat{c}^* = b^{(d+2-\eta)/2}. \quad (5.112)$$

Of course, other critical exponents must come from an analysis of the fixed point spectrum.

Naturally one cannot, in general, implement this momentum shell transformation exactly. Nor can one be necessarily assured of smoothness, locality and aptness. However, in the same way that the one-dimensional Ising model can be treated exactly by the decimation or a block spin renormalization group - thus constituting an analogue to the quantum mechanical "particle in the box" problem - so can Gaussian models, as described in Sec. 4.7, be treated exactly by the momentum shell transformation. One might, indeed, regard the Gaussian model as the "hydrogen atom" of critical phenomena: unfortunately, however, in itself it is of distinctly less direct physical relevance than the hydrogen atom. Even so, as we shall show in the next section, a solution of the Gaussian model via the momentum shell renormalization group provides a foundation on which can be built a systematic expansion procedure for solving more realistic and challenging models!

## 6. Dimensionality Expansions

On the face of it, the Wilson momentum space renormalization group seems to suffer from the same afflictions as the previously-described real-space renormalization groups; but it turns out to have an overwhelming advantage in that the unavoidable approximations can now be made in a systematic and controlled way. It has thus proved possible to make many useful, novel, and incisive calculations with the momentum shell integration technique and, even in rather low orders of approximation, quite accurate numerical results have been obtained.

As explained, the momentum shell renormalization group requires the use of continuous spins with, say,  $n$ -components so that  $\vec{s}_{\underline{x}} = (s_{\underline{x}}^{(\mu)})_{\mu=1,2,\dots,n}$  with  $-\infty < s^{(\mu)} < \infty$ . At first sight this precludes its application to discrete spin systems, like the spin 1/2 Ising model, or to systems with fixed length spins like the classical Heisenberg model with, say,  $|\vec{s}_{\underline{x}}| = 1$ . However, this view proves too naive since, via a Kac-Hubbard-Stratonovich transformation, such models can be transformed exactly into thermodynamically equivalent continuous spin models with definite spin weighting functions of the general sort discussed in Sec. 4.7. How this works is explained in Appendix A. Here we will assume that a continuous spin model is given and we start by transforming it into a Fourier space representation suitable for application of the momentum shell procedure.

### 6.1 Transformation of the Hamiltonian

Following (4.44) we consider the total reduced Hamiltonian expressed in real space variables as

$$\bar{\mathcal{H}} = - \mathcal{H}_{\text{int}}(\vec{s}_{\underline{x}})/k_B T - \sum_{\underline{x}} w(\vec{s}_{\underline{x}}), \quad (6.1)$$

where the interaction Hamiltonian (or, for true spins, the "exchange" Hamiltonian) is given by

$$\mathcal{H}_{\text{int}}(\vec{s}_{\underline{x}}) = -\frac{1}{2} \sum_{\underline{x}_1} \sum_{\underline{x}_2} J(\underline{x}_1 - \underline{x}_2) \vec{s}_{\underline{x}_1} \cdot \vec{s}_{\underline{x}_2}, \quad (6.2)$$

while the site vectors  $\underline{x}$  range over a  $d$ -dimensional hyper-cubic lattice of spacing  $\underline{a}$ . The single-spin weighting function is expanded as

$$w(\vec{s}) = \frac{1}{2} |\vec{s}|^2 + \tilde{u} |\vec{s}|^4 + \tilde{v} |\vec{s}|^6 + \dots \quad (6.3)$$

It is worth recalling at this point that, as mentioned in Sec. 4.7 and demonstrated in Appendix A, the discrete variable spin  $\frac{1}{2}$  Ising model and the classical, fixed-length,  $n$ -vector models can all be cast exactly in the form of continuous spin models as considered here with no approximation. Now we introduce momentum space variables  $\hat{s}_{\underline{g}}$  via (5.103) with the inverse relation (5.101). It is then straightforward to transform (6.1) into the form

$$\begin{aligned} \bar{\mathcal{H}} = & -\frac{1}{2} \frac{1}{N} \sum_{\mathbf{q}} [1 - \hat{K}(\mathbf{q})] \hat{s}_{\mathbf{q}} \cdot \hat{s}_{-\mathbf{q}} \\ & - \tilde{u} \frac{1}{N^3} \sum_{\mathbf{q}_1} \sum_{\mathbf{q}_2} \sum_{\mathbf{q}_3} (\hat{s}_{\mathbf{q}_1} \cdot \hat{s}_{\mathbf{q}_2}) (\hat{s}_{\mathbf{q}_3} \cdot \hat{s}_{\mathbf{q}_4}) + \dots, \end{aligned} \quad (6.4)$$

where the wave vectors appearing in the multiple sums are restricted by

$$\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4 = 0, \quad \mathbf{Q}, \quad (6.5)$$

for the fourth order term, and similarly in higher orders, where  $\mathbf{Q}$  is any reciprocal lattice vector. The interactions now appear via the Fourier transform

$$1 - \hat{K}(\mathbf{q}) = 1 - \sum_{\mathbf{x}} e^{i\mathbf{q} \cdot \mathbf{x}} \frac{J(\mathbf{x})}{k_B T}. \quad (6.6)$$

If the couplings are of reasonably short range it is possible to expand  $\hat{K}(\mathbf{q})$  in a power series in  $\mathbf{q}$  in which, for symmetry reasons the linear term vanishes and the quadratic term is proportional to  $|\mathbf{q}|^2$ . The result can be written in the form

$$1 - \hat{K}(\mathbf{q}) = \frac{T - T_0}{T} + \frac{\hat{J}(0)}{k_B T} R_0^2 q^2 + o(q^4), \quad (6.7)$$

where the mean field critical temperature,  $T_0$ , has been introduced via

$$k_B T_0 = \hat{J}(0) = \sum_{\mathbf{x}} J(\mathbf{x}), \quad (6.8)$$

while  $R_0$  measures the range of the interactions.

Of course we will be interested in the thermodynamic limit  $N \rightarrow \infty$ . The wavevector sums then become integrals and, to simplify formulae, we will employ the shorthand notation

$$\frac{a^{-d}}{N} \sum_{\mathbf{q}} \int \frac{d^d \mathbf{q}}{(2\pi)^d} \equiv \int_{\mathbf{q}}. \quad (6.9)$$

A rescaling of the spin variables by the substitution

$$\hat{s}_{\mathbf{q}} = \vec{\sigma}_{\mathbf{q}} \left(\frac{T}{T_0}\right)^{1/2} \frac{1}{R_0 a^{d/2}}, \quad (6.10)$$

transforms  $\bar{\mathcal{H}}$  into the more convenient standard form

$$\bar{\mathcal{H}} \approx -\frac{1}{2} \int_{\mathbf{q}} (r + eq^2) \vec{\sigma}_{\mathbf{q}} \cdot \vec{\sigma}_{-\mathbf{q}} \quad (6.11)$$

$$-u \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} \int_{\mathbf{q}_3} \sum_{\mu, \nu=1}^n \sigma_{\mathbf{q}_1}^{\mu} \sigma_{\mathbf{q}_2}^{\mu} \sigma_{\mathbf{q}_3}^{\nu} \sigma_{\mathbf{q}_4}^{\nu},$$

with

$$g_1 + g_2 + g_3 + g_4 = 0, \quad (6.12)$$

which is known as the Landau-Ginzburg-Wilson (LGW) reduced Hamiltonian or as a field-theoretic Hamiltonian (or "action"). The coefficient  $e$  has been introduced in the quadratic term but, at this stage, it is simply equal to unity. The leading coefficient,  $r$ , now stands in for the temperature since one has

$$r = \frac{T-T_0}{T_0 R_0^2} = \frac{t_0}{R_0^2}, \quad (6.13)$$

where  $t_0$  measures the deviation from mean field criticality. Finally, the coefficient of the fourth order term becomes

$$u = \left( \frac{\tilde{u}(T)}{T_0} \right)^2 \frac{a^{d-4}}{(R_0/a)^4}, \quad (6.14)$$

which reveals, for the first time, how the deviation in dimensionality

$$\varepsilon = 4 - d, \quad (6.15)$$

arises naturally. Note that "umklapp" processes, with  $Q \neq 0$ , have been ignored,  $q^4$ ,  $q^6$ , ... terms have been dropped and sixth and higher order terms in  $\sigma_{\mathbf{q}}$  have been neglected: in the end one can (and should) return to check that all of these contributions represent irrelevant variables in the domain of interest. As usual,  $\vec{\mathcal{H}}$  can be regarded as a point,  $(r, e, u, v, \dots)$  in the space of Hamiltonians where  $v$  represents the coefficient of the sixth order terms and so on.

## 6.2 Computing with continuous spins

Since we have continuous spin variables, computing the trace of the Boltzmann factor involves multiple integrals over the spins as indicated in (5.104) and (5.105). The general Hamiltonian may be written

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}_2 - u \bar{\mathcal{H}}_4 - v \bar{\mathcal{H}}_6 + \dots, \quad (6.16)$$

where the first term is quadratic in the  $s_{\mathbf{x}}$  (and, hence, in the  $\hat{s}_{\mathbf{q}}$  and  $\vec{\sigma}_{\mathbf{q}}$ ), while the second is quartic, and so on. If the higher order terms could be dropped, leaving only a quadratic or "free-field" Hamiltonian, the Boltzmann factor  $\exp(\bar{\mathcal{H}}_2)$  would decompose into a product of Gaussian functions of the individual  $\vec{\sigma}_{\mathbf{q}}$  variables. The trace integrations would then be trivial! The most obvious way of handling the higher order terms is thus to treat them as a

perturbation and to attempt an expansion of the free energy in powers of  $u$ ,  $v$ , etc. Confining ourselves to the quartic term we would then expand as

$$e^{\bar{\mathcal{H}}_2 + u\bar{\mathcal{H}}_4} = e^{\bar{\mathcal{H}}_2} \left( 1 - u\bar{\mathcal{H}}_4 + \frac{1}{2!} u^2 \bar{\mathcal{H}}_4^2 + \dots \right), \quad (6.17)$$

which represents a Gaussian function times polynomials in the  $\vec{\sigma}_q$ . One is now confronted by various combinations of products of integrals, all of the same basic type, namely,

$$\int_{-\infty}^{\infty} d\sigma_q^\mu e^{-\frac{1}{2}(r+eq^2)|\sigma_q^\mu|^2} |\sigma_q^\mu|^k = \frac{I_k}{(r+eq^2)^{(k+1)/2}}, \quad (6.18)$$

which integrate out as shown,  $I_k$  being a constant which vanishes for  $k$  odd.

Since the wavevectors,  $q$ , form a quasicontinuum, the products referred to become infinite products in the thermodynamic limit and the process of taking the logarithm of the overall trace to obtain an expression for the free energy, thus yields momentum integrals of the form

$$\mathcal{J}_k^{(d)} = \int \frac{d^d q}{(r+q^2)^{k/2}}. \quad (6.19)$$

If we proceed straight ahead in a perturbation theoretic spirit we now confront a major problem, namely, the so-called infra-red divergences. In the absence of the perturbation we have a Gaussian model which becomes critical as  $r \rightarrow 0$ . With the quartic term present we actually expect the critical point to be depressed to negative  $r$  but, in any case, if we want to study the critical region we must at least consider  $r \rightarrow 0$ . However, in that limit all the integrals will diverge for large enough  $k$  owing to the singularity of the integrand as  $q \rightarrow 0$  (whence the terminology "infra-red"). Specifically, counting powers of momentum shows that  $\mathcal{J}_k^{(d)}$  diverges as  $k \rightarrow 0$  whenever  $d < k$ . Since even the leading term in (6.17) involves  $k = 4$  we see that the naive perturbation method fails immediately for  $d < 4$ !

The Wilson approach circumvents this basic problem by never actually integrating over momenta beneath the reduced cutoff  $q_\Lambda/b$  (see Fig. 5.8). Thus no divergences are encountered.

### 6.3 Implementation of momentum shell renormalization

In order to implement Wilson's momentum shell renormalization group in a perturbative manner we split the original Hamiltonian (6.11) in the form

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}^< + \bar{\mathcal{H}}_2^> - u\bar{\mathcal{H}}_4^>, \quad (6.20)$$

where  $\bar{\mathcal{H}}^<$  includes all those parts of  $\bar{\mathcal{H}}$  which contain only spins,  $\bar{\mathcal{H}}_2^>$  is the Gaussian, free-field or quadratic part of the total Hamiltonian with spins  $\sigma_q^>$  with momenta lying in the outer shell (see Fig. 5.8). Finally  $u\bar{\mathcal{H}}_4$  consists of all the remaining terms which involve spins  $\sigma_q^>$  (as well as, in general, some  $\sigma_q^<$ ).

The trace operation

$$\text{Tr}_{N-N'}^{\sigma_q^>} \{ e^{\bar{\mathcal{H}}(\sigma_q^<, \sigma_q^>)} \}, \quad (6.21)$$

which we want to carry out, can be expressed conveniently in a perturbation series if we make use of the notation

$$\langle X \rangle = \frac{\text{Tr}_{N-N'}^{\sigma_q^>} \{ X e^{\bar{\mathcal{H}}_2^>} \}}{\text{Tr}_{N-N'}^{\sigma_q^>} \{ e^{\bar{\mathcal{H}}_2^>} \}}, \quad (6.22)$$

which represents averaging with the free-field Hamiltonian over only the higher momentum fluctuations, i.e., those with  $q$  in the shell. Now the renormalized Hamiltonian can be written

$$\bar{\mathcal{H}}' = \left[ \ln \left( \text{Tr}^{\sigma_q^>} \left\{ e^{\bar{\mathcal{H}}^<} e^{\bar{\mathcal{H}}_2^>} e^{-u\bar{\mathcal{H}}_4^>} \right\} \right) \right]_{\sigma_q^>} \Rightarrow \sigma_q^>, \quad (6.23)$$

Notice first that the factor  $e^{\bar{\mathcal{H}}^<}$  commutes with the trace operation since it involves only spins  $\sigma_q^<$ . Then, using the notation (6.22), one readily establishes the expansion

$$\begin{aligned} \bar{\mathcal{H}}' &= \left[ \bar{\mathcal{H}}^< + \ln \left( \text{Tr}^{\sigma_q^>} \left\{ e^{\bar{\mathcal{H}}_2^>} \right\} \right) \right. \\ &\quad \left. + \ln \left( 1 - u \langle \bar{\mathcal{H}}_4^> \rangle + \frac{1}{2} u^2 \langle (\bar{\mathcal{H}}_4^>)^2 \rangle + \dots \right) \right]_{\sigma_q^>} \Rightarrow \sigma_q^>, \\ &= \left[ \bar{\mathcal{H}}^< + \ln \left( \text{Tr}^{\sigma_q^>} \left\{ e^{\bar{\mathcal{H}}_2^>} \right\} \right) - u \langle \bar{\mathcal{H}}_4^> \rangle \right. \\ &\quad \left. + \frac{1}{2} u^2 \left( \langle (\bar{\mathcal{H}}_4^>)^2 \rangle - \langle \bar{\mathcal{H}}_4^> \rangle^2 \right) + O(u^3) \right]_{\sigma_q^>} \Rightarrow \sigma_q^>, \end{aligned} \quad (6.24)$$

which we have performed to order  $u^2$ . We may now set about calculating more explicitly the recursion relations

$$\begin{aligned}
r' &= \mathcal{R}_r(r, e, u, \dots), \\
e' &= \mathcal{R}_e(r, e, u, \dots), \\
u' &= \mathcal{R}_u(r, e, u, \dots),
\end{aligned} \tag{6.25}$$

to successive powers of  $u$ . (We will neglect the constant term since it cannot enter the recursion relations for  $r$ ,  $e$ ,  $u$ , etc. and cannot play a role in determining critical exponents. However, it would be needed for studying the full free energy.) The details of the derivation are presented more fully in Appendix B. In the lowest orders one finds just

$$r' = \hat{c}^2 b^{-d} [r + 2u (\quad) + O(u^2)], \tag{6.26}$$

$$e' = \hat{c}^2 b^{-d-2} [e + 2u (\quad) + O(u^2)], \tag{6.27}$$

$$u' = \hat{c}^4 b^{-3d} u [1 - \frac{1}{2}u (\quad) + O(u^2)]. \tag{6.28}$$

The origin of the factors  $\hat{c}$  here is easy to understand. The spin rescaling (5.110) introduces a factor of  $\hat{c}$  for each unrenormalized spin component  $\sigma_g^\mu$ , and so  $r'$  and  $e'$ , which are associated with the  $\sigma_g^\mu \sigma_{-g}^\mu$  terms in  $\overline{\mathcal{H}}'$ , acquire factors of  $\hat{c}^2$ . Likewise  $u$ , which is associated with the quartic term, acquires a factor  $\hat{c}^4$ . The factors of  $b$  come from the spatial rescaling (5.109). Since the momentum integrals in (6.11) transform as

$$\int_{\mathcal{g}} (r + eq^2) \times \dots \Rightarrow \int b^{-d} \frac{d^d q'}{(2\pi)^d} [r + e (q'/b)^2] \times \dots, \tag{6.29}$$

the expression for  $r'$  acquires a factor of  $b^{-d}$ , and that for  $e'$  a factor  $b^{-d-2}$ . Similarly, since there are integrations over three different momenta involved in the quartic spin term, a factor of  $b^{-3d}$  enters for  $u$ . The reader should check these statements carefully: although they involve only dimensional analysis they turn out to be a most crucial ingredient!

Now it is clear that the choice of an overall scale factor remains at our disposal [as used in writing (6.10)]. This freedom can be used to fix one of the parameters: following (6.11) we will choose to maintain the constraint (or normalization)

$$e' = e = 1. \tag{6.30}$$

The reason for this choice is that the  $q^2$  term (or, in real space, the gradient squared term) is the one that sets the physical length scales. It follows that the

spin rescaling factors are determined (to leading order) by

$$\hat{c}^2 = b^{d+2} \quad \text{and} \quad c^2 = \hat{c}^2/b^{2d} = b^{-d+2}. \quad (6.31)$$

If we find a nontrivial fixed point under these conditions it must, via (5.112) or (5.73), mean that

$$\eta = 0 + O(u). \quad (6.32)$$

The other two recursion relations then give us, in leading i.e., zeroth order, the results

$$r' = b^2 r, \quad (6.33)$$

$$u' = b^{4-d} u = b^\epsilon u. \quad (6.34)$$

The appearance of the factor  $b^\epsilon$  in this last equation is a vital feature. Note that it simply reflects the canonical dimensions of  $u$  (in terms of lengths) as revealed in (6.14).

#### 6.4 The Gaussian fixed point

The only fixed point that exists in the zeroth order approximation developed above is the Gaussian fixed point given by

$$r_G^* = 0, \quad u_G^* = 0. \quad (6.35)$$

With  $u = 0$  the recursion relations are now diagonal as they stand, and the Gaussian eigenvalues are evidently

$$\Lambda_1 \equiv \Lambda_r = b^2, \quad \text{so} \quad \lambda_1 = \lambda_r = 2, \quad (6.36)$$

and

$$\Lambda_u = b^\epsilon, \quad \text{so} \quad \lambda_u = \epsilon. \quad (6.37)$$

We see that the parameter  $u$  changes from being irrelevant at large  $d$  to being relevant at the border-line dimensionality  $\epsilon = 0$ , i.e.,  $d = 4$ . Thus for  $d > 4$  the quartic spin terms prove to be irrelevant, and  $u \rightarrow 0$  under renormalization. The first, relevant eigenvalue must clearly be identified as the thermal eigenvalue. Through (5.38) or (5.96) we thus find

$$v = \frac{1}{2}, \quad (d > 4), \quad (6.38)$$

which is the classical value! At  $d = 4$  we see that  $u$  is marginal and, to this order, does not shift under renormalization. Finally,  $u$  is relevant for  $d < 4$  and flows away from the Gaussian fixed point. The crucial question is: "Where to?"

Before answering that question, for which we must study the recursion relations in higher order, note that we may still consider the full scaling form (5.90) for the free energy around the Gaussian fixed point. Evidently, crossover away from Gaussian) critical behavior (for which all other exponents also prove to be classical) is controlled by the scaled combination

$$u/r^{\phi_u} \quad \text{with } \phi_u = \frac{\lambda_u}{\lambda_1} = \frac{1}{2} \varepsilon. \quad (6.39)$$

But notice now, from (6.13) and (6.14), that both  $r$  and  $u$  depend inversely on the range of the forces  $R_0$ . It follows that the range enters in the combination  $a^d/R_0^d t_0^{\varepsilon/2}$ . Hence, if  $R_0/a$  is large the Gaussian fixed point should describe the critical point (which will then look classical) until  $R_0 t_0^{\varepsilon/2d}/a$  becomes small. Since the exponent here is comparatively small (being  $1/6$  for  $d = 3$ ) the crossover to nonclassical behavior may take place rather slowly. The exponent we have found<sup>23</sup> for the long range crossover agrees with that following from the Ginzburg criterion for the validity of classical theory.<sup>24-27</sup> Its small value serves to explain, for example, why the BCS theory of superconductivity, which is a classical or mean field theory, works so well in practice; the ratio  $R_0/a$  is there measured by  $T_F/T_c$ , where  $T_F$  is the Fermi temperature and  $T_c$  is the superconducting transition temperature. (Of course the BCS theory is quantum-mechanical in nature: the word "classical" here, as elsewhere, refers only to the neglect of fluctuations in the statistical mechanical treatment.)

### 6.5 The renormalization group to order $\varepsilon$

If we are to obtain useful results for  $d < 4$ , the unstable flow from the Gaussian fixed point must, for at least one direction of flow, terminate at some new nontrivial fixed point. Since (6.34), the zeroth order recursion relation for  $u$ , is linear this is impossible unless we carry the perturbation calculation explicitly to at least the next order. To do this, a diagrammatic formulation, modeled on field theory, is helpful as sketched in Appendix B: the requisite analysis serves to fill the blanks in (6.26) to (6.28) and yields,

$$r' = \hat{c}^2 b^{-d} [r + 4u (n+2) \int_{\mathbf{q}}^{\mathbf{q}'} \frac{1}{r+q^2} + O(u^2)], \quad (6.40)$$

$$e' = \hat{c}^2 b^{-d-2} [1 + 0 + 0 (u^2)], \quad (6.41)$$

and

$$u' = \hat{c}^4 b^{-3d} u \left[ 1 - \frac{1}{2} u \cdot 8(n+8) \int_g^> \frac{1}{(r+q^2)^2} + 0 (u^2) \right]. \quad (6.42)$$

We see from the second relation that in order to maintain  $e' = e = 1$ , we must set

$$\hat{c}^2 = b^{d+2+0(u^2)}, \quad (6.43)$$

This implies that if there is a nontrivial fixed point at some  $u = u^* \neq 0$ , then the critical point decay exponent satisfies

$$\eta = 0(u^{*2}). \quad (6.44)$$

Although this is not an explicit formula it gives us some understanding of why  $\eta$  is so small in most physical systems relative to the deviations of other exponents from their classical values.

Now if we substitute with (6.43) in (6.42) the prefactor  $b^\epsilon$  appears again. Let us, then, invoke the idea of continuous dimensionality and enquire as to what happens if  $\epsilon = 4 - d$  is small! First we can write

$$b^\epsilon = 1 + \epsilon \ln b + 0(\epsilon^2). \quad (6.45)$$

Then, keeping only terms linear in  $\epsilon$  and  $u$ , the recursion relation for  $u$  can be rewritten as

$$u' - u = u \left[ \epsilon \ln b - 4 (n+8) u \int_g^> \frac{1}{(r+q^2)^2} \right], \quad (6.46)$$

where the integral is now to be evaluated at  $d = 4$ . Evidently this recursion relation has a new fixed point at  $u = u^* \propto \epsilon / (n+8)$ . Since, by supposition,  $\epsilon$  is small,  $u^*$  is also small and therefore we can neglect the  $0(u^2)$  corrections since, in the neighborhood of this fixed point they will be of order  $\epsilon^2$ . This is the crux of the  $\epsilon$ -expansion idea: by expanding in powers of  $\epsilon$  we may utilize the field-theoretic perturbation theory in powers of  $u$  in a systematic way. We rely on the renormalization group framework since although the initial, physical value of  $u$  may well be 'large', i.e., of order unity, the flow of the critical trajectory to the fixed point allows us to calculate only for  $u$  small, of order  $\epsilon$ .

Clearly the integrals require a little further thought. Note, first, that if we approximate the momentum shell by a hypersphere (instead of using a hypercube) we

have

$$\int_{\mathcal{G}}^{\>} \equiv \int_{q_{\Lambda}/b}^{q_{\Lambda}} \hat{C}_d q^{d-1} dq, \quad (6.47)$$

where the area of the unit hypersphere is, as before,

$$C_d = (2\pi)^d \hat{C}_d = 2\pi^{d/2} / \Gamma(\frac{1}{2}d). \quad (6.48)$$

Since the gamma function is an analytic function of its argument, we certainly can see that any spherically symmetric integrals, such as involved here, can be extended to continuous dimension. (As indicated in Appendix C one can even extend hypercubic lattices to continuous  $d$ .) Accepting the hyperspherical approximation we can perform both the needed integrals in the critical region, i.e., for small  $r$ . One finds

$$\begin{aligned} \int_{\mathcal{G}}^{\>} \frac{1}{(r+q^2)^2} &= \int_{q_{\Lambda}/b}^{q_{\Lambda}} \frac{1}{8\pi^2} \frac{q^3 dq}{q^4} + O(\epsilon, r), \\ &= K_2 \ln b + O(\epsilon, r), \end{aligned} \quad (6.49)$$

with  $K_2 = 1/8\pi^2$  and, similarly,

$$\int_{\mathcal{G}}^{\>} \frac{1}{(r+q^2)} = K_1(1-b^{-2}) - K_2 r \ln b + O(\epsilon, r^2). \quad (6.50)$$

The constant  $K_1$ , has the value  $q_{\Lambda}^2/16\pi^2$  in the spherical approximation but its actual value proves to be immaterial as regards all universal quantities; conversely the value of  $K_2$  in these two equations is independent of the approximation, as is easily seen by more careful analysis.

We can now write the recursion relations correct to relative order  $\epsilon$ ,  $u$  and  $ur$  as

$$r' = b^2 r [1 - 4(n+2)K_2 u \ln b] + 4(n+2)K_1 (b^2 - 1)u, \quad (6.51)$$

$$u' = u + u \ln b [\epsilon - 4(n+8)K_2 u]. \quad (6.52)$$

Before analyzing these relations it is worth noting that they can be cast in differential form, as in (5.75) to (5.77), by putting

$$b = e^{\delta\ell} = 1 + \delta\ell + O(\delta\ell^2), \quad (6.53)$$

and taking the limit  $\delta\ell \rightarrow 0$ . Thus one obtains

$$\frac{dr}{d\ell} = [2-4(n+2)K_2]r + 8(n+2)K_1 u, \quad (6.54)$$

$$\frac{du}{d\ell} = u[\varepsilon - 4(n+8)K_2 u]. \quad (6.55)$$

### 6.6 The n-vector fixed point

Now we can investigate the new fixed point.<sup>28</sup> From (6.52) or (6.55) we have

$$u^* = \frac{\varepsilon}{4(n+8)K_2} = \frac{2\pi^2 \varepsilon}{(n+8)}, \quad (6.56)$$

which is, of course, only correct to order  $\varepsilon$ . Then from (6.51) or (6.54) we find

$$r^* = -\frac{K_1}{K_2} \left( \frac{n+2}{n+8} \right) \varepsilon, \quad (6.57)$$

where the ratio  $K_1/K_2$  has the nonuniversal value  $\frac{1}{2} q_\Lambda^2$  in the spherical approximation. Evidently this new fixed point "breaks off" from the Gaussian fixed point as the dimensionality falls below the borderline  $d = 4$ . (Actually, it exists also for  $d > 4$  but at negative  $u^*$  where it is unstable and plays no role since  $u > 0$  is needed if the partition function is to be well-defined, at least in the absence of any higher order stabilizing terms.) It is easy to determine the flows in the  $(r, u)$  plane near the fixed points. Their appearance for  $\varepsilon > 0$ , i.e.,  $d < 4$  is shown in Fig. 6.1.

To determine the critical exponents and test the new fixed point for its stability we must linearize about  $(r^*, u^*)$ . If we work with the discrete recursion relations (6.51) and (6.52), we may write

$$\Delta r = r - r^* \quad \text{and} \quad \Delta u = u - u^*, \quad (6.58)$$

and so obtain the matrix form

$$\begin{bmatrix} \Delta r \\ \Delta u \end{bmatrix} = \begin{bmatrix} b^2(1 - \frac{n+2}{n+8} \varepsilon \ln b) & 4(n+2)K_1(b^2-1) \\ 0 & 1 - \varepsilon \ln b \end{bmatrix} \begin{bmatrix} \Delta r \\ \Delta u \end{bmatrix}. \quad (6.59)$$

Note that  $K_2$  has cancelled out! The eigenvalues follows at once as

$$\Lambda_1 \approx b^2 \left( 1 - \frac{n+2}{n+8} \varepsilon \ln b \right) \approx b^{2 + \frac{n+2}{n+8} \varepsilon}, \quad (6.60)$$

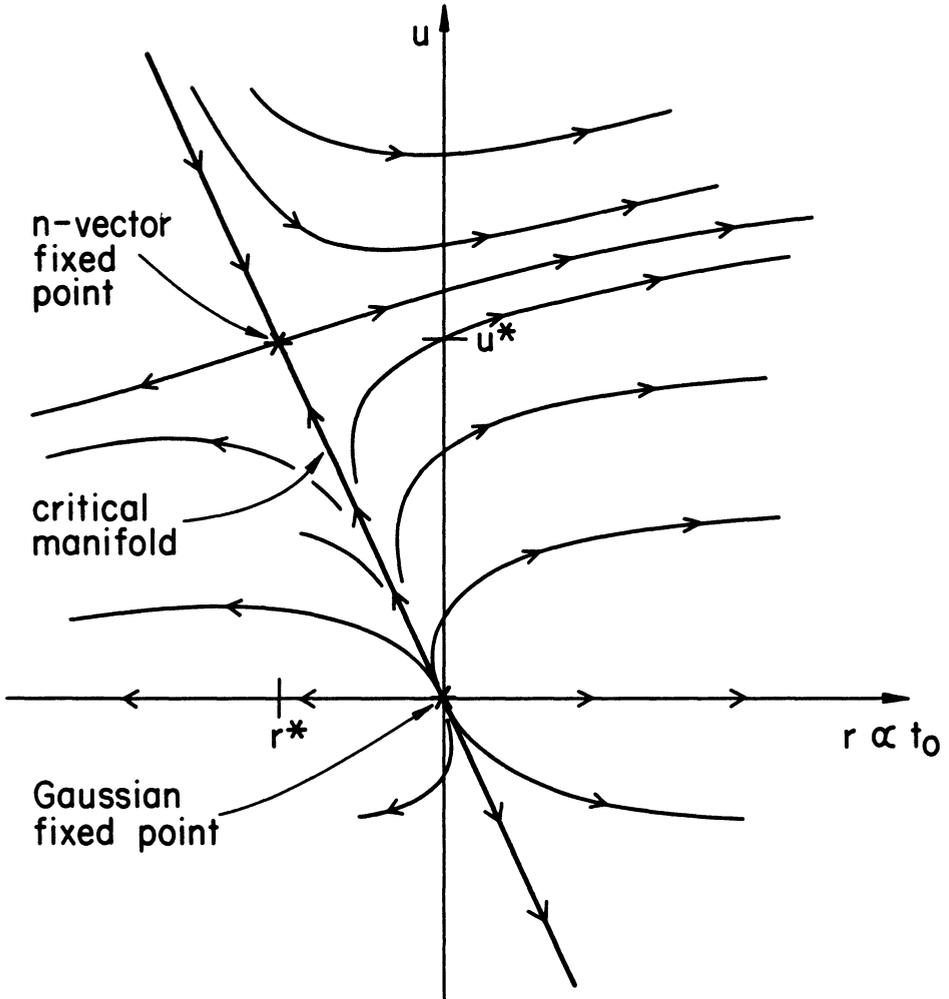


Fig. 6.1 Sketch of the renormalization group flows in the  $(r, u)$  plane for small  $\epsilon = 4 - d$ . Note that the critical manifold (or trajectory) is straight only to order  $\epsilon$ .

$$\Lambda_u \approx 1 - \epsilon \ln b \approx b^{-\epsilon}, \quad (6.61)$$

so that we have

$$\lambda_1 \equiv \lambda_\xi = 2 - \frac{n+2}{n+8} \epsilon + O(\epsilon^2), \quad (6.62)$$

$$\lambda_u = -\epsilon + O(\epsilon^2). \quad (6.63)$$

From the last result we see that  $u$  represents an irrelevant variable about the new fixed point when  $\epsilon > 0$ : in other words this fixed point is "stable" when  $d < 4$  and hence "controls" the flow in place of the Gaussian fixed point (see Fig. 6.1) which, as we found, is now unstable. [Strictly we should say "stable (or unstable) on the critical manifold" but the restriction is left unstated in practical terminology since the expected, relevant, unstable directions are always understood.]

Finally, we may use the renormalization group eigenvalues to compute the critical exponents. Thus for the correlation length we find from (5.96)

$$\nu = \frac{1}{\lambda_1} = \frac{1}{2} + \frac{n+2}{4(n+8)} \epsilon + O(\epsilon^2). \quad (6.64)$$

If we recall our result for the correlation decay exponent, namely,

$$\eta = O(u^{*2}) = 0 + O(\epsilon^2), \quad (6.65)$$

we may use a scaling relation (which may be verified independently by more detailed calculations) to find

$$\begin{aligned} \gamma &= (2-\eta)\nu = [2 + O(\epsilon^2)]/\lambda_1, \\ &= 1 + \frac{n+2}{2(n+8)} \epsilon + O(\epsilon^2). \end{aligned} \quad (6.66)$$

Likewise the hyperscaling relations yield  $\alpha$  from  $2 - \alpha = d\nu$  and

$$\begin{aligned} \beta &= \frac{1}{2}(d-2+\eta)\nu \\ &= \frac{1}{2} - \frac{3}{2(n+8)} \epsilon + O(\epsilon^2). \end{aligned} \quad (6.67)$$

The other thermodynamic exponents follow similarly. In addition we obtain something new, namely, the leading correction-to-scaling exponent [see (5.93) et seq.] which is associated with  $u$  and hence given by

$$\theta = -\phi_u = -\frac{\lambda_u}{\lambda_1} = \frac{1}{2} \epsilon + O(\epsilon^2). \quad (6.68)$$

Historically, the first theoretical predictions for the value of  $\theta$  come from the renormalization group  $\epsilon$  expansion.

### 6.7 Some numerics

It is natural to enquire how well, if at all, the  $\epsilon$  expansion works! To answer, recall first that the limit  $n \rightarrow \infty$  should reproduce the spherical model (see Sec. 4.6). The exact results for the spherical model include  $\beta = 1/2$  which certainly agrees with (6.67) and, more interestingly,

$$\gamma = \frac{2}{d-2} = 1 + \frac{1}{2} \epsilon + \frac{1}{4} \epsilon^2 + \frac{1}{8} \epsilon^3 + \dots, \quad (6.69)$$

which confirms (6.66) precisely! In this limit we see, in fact, that the  $\epsilon$  expansions for the exponents represent convergent power series with a radius of convergence  $\epsilon_c = 2$ .

For finite  $n$  it seems more likely that the  $\epsilon$  expansion is only asymptotic (but, probably, "Borrel summable"). Nevertheless, we may, optimistically, hope that  $\epsilon = 1$  (for  $d = 3$ ) is relatively "small" in that it is only halfway to the undoubted breakdown around  $\epsilon_c = 2$ . This optimism turns out to be surprisingly well justified. Indeed, even the first order expansions yield values in much better agreement with bulk ( $d=3$ )-dimensional experiments than does classical theory. Thus from (6.66) and (6.67) we find

$$\begin{aligned} \gamma(d=3, n=1) &\approx 1\frac{1}{6} \approx 1.67, & \text{while } \gamma_{\text{expt}} &\approx 1.24, \\ \beta(d=3, n=1) &\approx \frac{1}{3}, & \text{while } \beta_{\text{expt}} &\approx 0.32 - 0.33. \end{aligned}$$

In second order the results are even more encouraging. Thus from

$$\alpha = \frac{4-n}{2(n+8)} \epsilon - \frac{(n+2)^2(n+28)}{4(n+8)^2} \epsilon^2 + O(\epsilon^3), \quad (6.70)$$

one obtains

$$\text{for } \begin{array}{l} \alpha(d=3) \approx 0.08, \quad -0.02, \quad \text{and} \quad -0.10, \\ n = 1, \quad 2, \quad \text{and} \quad 3, \end{array}$$

respectively, which correlates well with the observed values

$$\alpha_{\text{expt}} \approx 0.11, \quad -0.02, \quad \text{and} \quad -0.14.$$

The trends with  $n$  are clearly reproduced and the divergence of the specific heat for  $n > 2$ , but not for  $n < 2$ , is also predicted.

By working harder one can calculate further terms in the expansion. Thus correct to fourth order one knows<sup>29</sup>

$$\begin{aligned} \eta = & \frac{(n+2)\epsilon^2}{2(n+8)^2} \left\{ 1 + \frac{(-n^2+56n+272)}{4(n+8)^2} \epsilon \right. \\ & + \left[ \frac{(-5n^4-230n^2+1124n^2+17920n+46144)}{16(n+8)^2} - \frac{24\zeta(3)(5n+22)}{(n+8)} \right] \epsilon^2 \\ & \left. + O(\epsilon^3) \right\}, \end{aligned} \quad (6.71)$$

which is interesting because of the appearance of the Riemann zeta function  $\zeta(s) = \sum_1^\infty n^{-s}$  in fourth order. All the other exponents are now also known to this order.<sup>30</sup> At this stage, however, one does not obtain better numerical results if one merely truncates the expansion: however, with suitable methods of summation<sup>30</sup> rather satisfactory results are obtained which, for the most part, appear to be accurate to within two or three parts in the third decimal place!

### 6.8 Further developments in brief

Having explained the concepts of scaling and universality, and having laid the foundations of renormalization group theory, these lectures must end. In a more extended course we would, at this point, proceed to survey some of the many significant applications of renormalization group theory. First we might demonstrate the appearance of factors like  $\ln t$  to special powers at the borderline dimensionality. Then, following the historical developments, we might consider the effect of long range forces<sup>31</sup> with a spin coupling decaying as

$$J(\underline{R}) \sim 1/R^{d+\sigma} \quad \text{as } R \rightarrow \infty, \quad (6.72)$$

with  $\sigma > 0$ . These are somewhat artificial but for  $\sigma < 2$  we would discover new critical behavior with a new borderline dimensionality at  $d = 2\sigma < 4$  about which we could construct a modified  $\epsilon$ -expansion in powers of  $\epsilon_\sigma = 2\sigma - d$ .

Next it would be logical to examine the crossover from, say, Heisenberg ( $n=3$ ) critical behavior to Ising ( $n=1$ ) or XY ( $n=2$ ) behavior induced by anisotropy in the spin-spin couplings and, hence, in the quadratic part of the LGW Hamiltonian. This could lead to a discussion of bicritical points as observed in many antiferromagnets.<sup>32</sup> Spin anisotropies of higher symmetry (induced physically by coupling to the lattice) in particular those entering as cubic symmetry breaking terms in the

quartic spin terms, can lead to quite new sorts of critical behavior: so also do long-range forces of dipole-dipole character which must be examined for real ferromagnets,<sup>33</sup> although for  $n > 2$  one finds that the numerical values of the exponents hardly change! For uniaxial, Ising-like dipolar ferromagnets, on the other hand, a striking new phenomenon occurs: the critical fluctuations at low momenta are suppressed and the borderline dimensionality drops from  $d=4$  to  $d=3$ .<sup>34,35</sup> Thus, except for subtle logarithmic correction factors, classical theory becomes correct for a real bulk system!

A borderline dimensionality  $d=3$  arises also in the description of tricritical points,<sup>36</sup> which is accomplished within a single-component ( $n=1$ ) model by allowing the coefficients of the  $s^4$  term to become negative but, as is needed for stability, retaining a term  $-vs^6$  with  $v > 0$ . Tricritical points are observed in multi-component fluid mixtures, in antiferromagnets, in superfluid helium three-four mixtures, etc.

Then one would want to describe the expansions in powers of  $1/n$  developed by Abe<sup>37</sup> (without explicit reference to the renormalization group) and those about the lower borderline dimensionality, putting  $d=2+\epsilon$ , devised to Polyakov.<sup>38</sup> The Kosterlitz-Thouless<sup>39-41</sup> theory of XY-like or  $n=2$  systems at the borderline dimensionality  $d=2$ , which describes thin superfluid helium films, and the subsequent Halperin-Nelson-Young<sup>42,43</sup> theories of two-dimensional melting would be tempting topics -- and so on! Even then, we would not have touched on the development and application of real space renormalization group techniques<sup>44</sup> including the versatile approximate renormalization group scheme of Migdal<sup>45</sup> and Kadanoff.<sup>46</sup> No mention would have been made of dynamical critical phenomena<sup>47</sup> and the application of renormalization groups in that context. Nor would we have discussed methods for calculating equations of state, or correlation functions, or crossover scaling functions,<sup>48,49</sup> or have described Wilson's method for solving the Kondo problem,<sup>50</sup> or applications to polymers, to liquid crystals, and more.

It is evident that to present even a sketchy account of all these topics would require much more time and space. Indeed, our task would grow to resemble that of giving a full account of the applications of quantum mechanics! Truly the renormalization group approach, and the associated ideas of scaling and universality, have become basic tools of the condensed matter theorist and are constantly being applied to new and more challenging problems. Happily, however, for the reader who wishes to enquire further there are now a selection of reviews and text books at various levels in which to browse and dig deeper. Some of these have already been mentioned in passing but for convenience these and a few more have been gathered together in the Bibliography. Note that the Bibliography makes no claim to completeness: indeed, we should add that in these lectures we have been somewhat cavalier in mentioning individual scientists and in making reference to the original literature. Accordingly, apologies are offered here to any who feel

unjustifiably unmentioned or otherwise slighted. The reader, however, should have no difficulty in entering the literature through the sources cited in the Bibliography: please do so with best wishes for stimulating study and fruitful discovery!

#### ACKNOWLEDGMENTS

The author is grateful to the CSIR of South Africa and Dr. G. Heymann for the invitation to lecture at the Advanced Course on Critical Phenomena at Stellenbosch. The hospitality of the Merensky Institute for Physics and its faculty, especially Professor Fritz J. W. Hahne and Professor Chris A. Engelbrecht, is much appreciated. The author is particularly indebted to Dr. Arthur G. Every for his expert and patient assistance in preparing a draft manuscript of the lectures and for seeing the full text through to final form. The ongoing support of the National Science Foundation, in part through the Materials Science Center at Cornell University, is gratefully acknowledged. Friends, students and colleagues too numerous to mention have helped the author learn and understand much of what is retold here: special thanks, however, are due to Cyril Domb, David M. Jasnow, Leo P. Kadanoff, Martin F. Sykes, Ben Widom and Kenneth G. Wilson.

A.G.E. gratefully acknowledges the hospitality extended to him at the University of Illinois and the CSIR for providing a travel grant. Thanks also to 'Tish' Watts for typing the manuscript.

APPENDIX A The Kac-Hubbard-Stratonovich Transformation

The Kac-Hubbard-Stratonovich transformation is a way of turning one model into another. This device has played an increasingly valuable role in the theory of critical phenomena. The main theoretical factor suggesting that various different models might be expected to transform into one-another is that of universality. To belong to the same universality class different models must, somehow, be mathematically equivalent, at least in their critical regions, even though they have quite different physical interpretations and contrasting mathematical formulations.

To illustrate how the transformation is carried out we will consider the simplest example, namely, a spin  $\frac{1}{2}$  Ising model; however, the approach to be described can be readily extended. (The interested reader should work through the case of the fixed-length n-vector model.)

Consider the general Ising model partition function

$$Z_N(K) = \text{Tr}_N^\sigma \left\{ \exp \left[ \sum_{(i,j)} K_{ij} \sigma_i \sigma_j \right] \right\}, \quad (\text{A1})$$

where the interactions satisfy

$$K_{ij} = J_{ij}/k_B T = K_{ji}, \quad (\text{with } K_{ii} \equiv 0), \quad (\text{A2})$$

and the sum in (A1) runs over all distinct pairs (i,j). The Ising spins  $\sigma_i$  in this expression can take on only the two values  $\pm 1$ , and so the operation of taking the trace over any spin means

$$\text{Tr}^\sigma = \frac{1}{2} \sum_{\sigma=\pm 1}. \quad (\text{A3})$$

The factor of 1/2 is incorporated in order to normalize the trace, i.e., so that  $\text{Tr}^\sigma \{1\} = 1$ . The aim now is to turn (A1) into a form which looks somewhat similar but involves a new set of continuous spin variables,  $s_i$  instead of the discrete  $\sigma_i$ . The result which we will obtain is

$$Z_N(K) = e^{f_0(K)} \int_{-\infty}^{\infty} \prod_{i=1}^N ds_i \exp \left[ - \sum_{(i,j)} O_{ij} s_i s_j \right] \exp \left[ - \sum_i w(s_i) \right], \quad (\text{A4})$$

where  $f_0(K)$  is a smooth, analytic function of  $K$  and just provides a background free energy with (in general) no interesting critical behavior. The integrations are performed over all the continuous spin variables,  $s_i$ , of which there are as many as there were original Ising spins; the limits for each integration are  $-\infty$  and  $+\infty$ . The spin weighting function comes out to be

$$-w(s_i) = \ln(\cosh s_i) - \frac{1}{2} O_{ii} s_i^2 = -\frac{1}{2} (O_{ii} - 1) s_i^2 - \frac{1}{12} s_i^4 + \dots \quad (\text{A5})$$

A weighting function is, of course, necessary because without it the integrals would diverge: in fact,  $Q_{ij}$ , which will be defined below, must be positive and sufficiently large. Our object now is to derive equations relating the original interactions  $K_{ij}$  to the new spin-spin interactions,  $-Q_{ij}$ , and provide a justification for the particular form assumed by the weighting function  $w(s_i)$ . (The fixed-length  $n$ -vector model yields a different but qualitatively similar set of weighting functions depending on  $n$ .)

A general reason for going over to continuous spin variables is that they are easier to deal with mathematically. In particular, in order to carry out the spatial Fourier transformations on the set of spin variables, which play such a vital role in the renormalization group  $\epsilon$ -expansion theory, the spins have to be continuous, unbounded variables.

For simplicity of exposition we will consider the fully ferromagnetic case  $K_{ij} = K_{ji} > 0$  ( $K_{ii} \equiv 0$ ). If we then make use of the inequality

$$\frac{1}{2} (\sigma_i + \sigma_j)^2 = \frac{1}{2} \sigma_i^2 + \frac{1}{2} \sigma_j^2 + \sigma_i \sigma_j = 1 + \sigma_i \sigma_j > 0, \quad (A6)$$

we can rewrite the interactions as

$$\sum_{(i,j)} K_{ij} \sigma_i \sigma_j = -\frac{1}{2} NP_0 + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N P_{ij} \sigma_i \sigma_j, \quad (A7)$$

where the symmetric,  $N \times N$  matrix

$$\underline{P} \equiv [P_{ij}] = P_0 \underline{I} + \underline{K} = [P_0 \delta_{ij} + K_{ij}], \quad (A8)$$

will be positive definite if  $P_0$  is chosen positive and sufficiently large: specifically, by (A6) it suffices to choose

$$P_0 > \max_i \left[ \sum_{\ell} K_{i\ell} \right]. \quad (A9)$$

If  $\underline{\sigma}^T = [\sigma_1, \dots, \sigma_N]$  is a row vector and  $\underline{\sigma}$  the corresponding (transposed) column vector, we can thus rewrite the partition function as

$$Z_N(\underline{K}) = e^{-\frac{1}{2} NP_0} \text{Tr}_N^{\underline{\sigma}} \left\{ e^{\frac{1}{2} \underline{\sigma}^T \underline{P} \underline{\sigma}} \right\}. \quad (A10)$$

Now consider another quadratic form in  $N$  continuous variables  $y_i$ , namely,

$$Q(\underline{y}) = \sum_{i=1}^N \sum_{j=1}^N Q_{ij} y_i y_j, \quad (A11)$$

where  $\underline{Q} = [Q_{ij}]$  is a symmetric, positive definite matrix: As such,  $\underline{Q}$  may be diagonalized by an orthogonal transformation with matrix  $\underline{Q}$ . Explicitly, in terms of

the new variables

$$\underline{x} = \underline{Q} \underline{y}, \quad (\text{A12})$$

the quadratic form becomes

$$Q(\underline{y}) = \sum_{r=1}^N \lambda_r x_r^2, \quad (\text{A13})$$

where the  $\lambda_r$  ( $r=1, \dots, N$ ) are the real, positive eigenvalues of  $\underline{Q}$ . It follows similarly that the determinant of  $\underline{Q}$  is given by

$$|\underline{Q}| = \prod_{r=1}^N \lambda_r. \quad (\text{A14})$$

The partition-function-like expression

$$I(\underline{Q}) = \int_{-\infty}^{\infty} d^N y e^{-\frac{1}{2} \underline{y}^T \underline{Q} \underline{y}}, \quad (\text{A15})$$

may now be evaluated by changing variables from  $\underline{y}$  to  $\underline{x}$  and noticing that the Jacobian of the transformation is +1 since  $\underline{Q}$  is orthogonal. The integrals are then just Gaussian and we obtain

$$I(\underline{Q}) = \prod_{r=1}^N \int_{-\infty}^{\infty} dx_r e^{-\frac{1}{2} \lambda_r x_r^2} = \prod_{r=1}^N \sqrt{\frac{(2\pi)}{\lambda_r}} = \frac{(2\pi)^{N/2}}{\sqrt{|\underline{Q}|}}. \quad (\text{A16})$$

Next let us make a simple shift in the variables  $\underline{y}$  according to

$$\underline{y} = \underline{s} + \underline{Q}^{-1} \underline{g} \quad \text{with} \quad \underline{y}^T = \underline{s}^T + \underline{g}^T \underline{Q}^{-1}, \quad (\text{A17})$$

where  $\underline{s}$  represents a new set of variables, which will eventually be identified as the continuous spin variables, while the  $\underline{g}$  here represent only fixed shift parameters! In terms of the new variables the quadratic form  $Q$  becomes

$$Q(\underline{y}) = \underline{s}^T \underline{Q} \underline{s} + \underline{s}^T \underline{g} + \underline{g}^T \underline{s} + \underline{g}^T \underline{Q}^{-1} \underline{g} = \underline{s}^T \underline{Q} \underline{s} + 2 \sum_{i=1}^N s_i \sigma_i + \underline{g}^T \underline{Q}^{-1} \underline{g}. \quad (\text{A18})$$

If we now choose the matrix  $\underline{Q}$  so that

$$\underline{Q} = \underline{P}^{-1} = 1/(\underline{P} \underline{Q} \underline{I} + \underline{K}), \quad (\text{A19})$$

we see quickly how (A15) and (A16) apply to the problem in hand since we obtain the identity

$$I(\underline{Q}) = e^{-\frac{1}{2} \underline{g}^T \underline{P} \underline{g}} \int_{-\infty}^{\infty} d^N s \exp \left[ -\frac{1}{2} \underline{s}^T \underline{Q} \underline{s} - \sum_{i=1}^N s_i \sigma_i \right],$$

$$\text{i.e., } I(Q) = \frac{(2\pi)^{N/2}}{|Q|} = (2\pi)^{N/2} \sqrt{|P|}. \quad (\text{A20})$$

Note that in the first line, the spin variables  $\sigma_i$  appear like nonuniform or random external fields acting upon the  $s$  variables. If we solve for the factor involving  $\tilde{P}$  and substitute in (A10) we can write the result as

$$Z_N(K) = \frac{e^{-\frac{1}{2}NP_0}}{(2\pi)^{N/2} \sqrt{|P|}} \text{Tr}_N^\sigma \left\{ \int_{-\infty}^{\infty} d^N s \exp \left[ - \sum_{(i,j)} Q_{ij} s_i s_j - \frac{1}{2} \sum_i Q_{ii} s_i^2 \right] e^{-\sum_i s_i \sigma_i} \right\}. \quad (\text{A21})$$

Note that the diagonal terms of the expression  $\frac{1}{2} \tilde{s}^T Q \tilde{s}$  have been separated off while what remains has been written as a sum over pairs of spins  $(i,j)$ .

At this point we make the crucial observation that the trace operation on the  $\sigma_i$  commutes with the integration over the  $s_i$  and affects only the last exponential factor in (A21). To perform the trace we use the simple result

$$\text{Tr}_{\sigma_i} \{ e^{-s_i \sigma_i} \} = \frac{1}{2} (e^{-s_i} + e^{+s_i}) = \cosh s_i, \quad (\text{A22})$$

so that, finally, we obtain

$$Z_N(K) = \frac{e^{-\frac{1}{2}NP_0}}{(2\pi)^{N/2} \sqrt{|P|}} \int_{-\infty}^{\infty} d^N s \exp \left[ - \sum_{(i,j)} Q_{ij} s_i s_j - \sum_i \left( \frac{1}{2} Q_{ii} s_i^2 - \ln \cosh s_i \right) \right]. \quad (\text{A23})$$

This is clearly in the anticipated form (A4) with the spin weighting function given by (A5) while the new interactions,  $-Q_{ij}$ , etc. follow from (A19) and (A8). Note that because  $\tilde{P}$  is positive definite by construction, so is  $Q$ ; it then follows, since  $\ln \cosh s$  varies only as  $|s|$  for large  $s$ , that the  $Q_{ij}$  coefficients are positive and sufficiently large to ensure convergence of the integrals over the  $s_i$ .

We have thus achieved an exact transformation of the discrete spin Ising model with couplings,  $K_{ij}$ , into a continuous spin model with new couplings,  $-Q_{ij}$ , defined via the inverse matrix  $(P_0 \tilde{I} + K)^{-1}$ . It is clearly of interest to gain some idea of the nature of these new interactions. To that end, let us suppose that  $K$  describes only nearest neighbor ferromagnetic couplings. A little thought then shows that  $K^2$  describes next-nearest neighbor couplings (plus some self-coupling), that  $K^3$  describes third-neighbor couplings (plus some further first neighbor couplings) and so on. Thus the identity

$$-P_0^2 Q = -P_0 \tilde{I} + K - P_0^{-1} K^2 + P_0^{-2} K^3 - \dots, \quad (\text{A24})$$

which is valid when  $P_0$  satisfies (A9), shows that the couplings  $-Q_{ij}$  are, in first approximation, the same as the  $K_{ij}$  but scaled by a factor  $P_0^{-2}$ . In higher approximation second-neighbor antiferromagnetic couplings appear but they are weaker by a factor  $1/P_0$ , and so on. Thus the new couplings are no longer of pure nearest-

neighbor character: however, they are of 'short range' in the sense that they decay exponentially with distance (and, evidently, they are oscillatory in sign although predominantly ferromagnetic in effect).

If the lattice is translationally invariant it is advantageous, as seen in Sec. 6.1, etc., to transform to continuous Fourier space spin variables,  $\hat{s}_{\mathbf{q}}$ . The couplings are then directly expressed, as in (6.4), in terms of the Fourier transform

$$\hat{Q}(\mathbf{q}) = \sum_{\mathbf{x}} e^{i\mathbf{q} \cdot \mathbf{x}} Q(\mathbf{x}) \quad \text{with} \quad Q(\mathbf{x}_i - \mathbf{x}_j) \equiv Q_{ij}. \quad (\text{A25})$$

This in turn is related to the corresponding transform  $\hat{K}(\mathbf{q})$  of  $K_{ij} \equiv K(\mathbf{x}_i - \mathbf{x}_j)$  through

$$-\hat{Q}(\mathbf{q}) = \frac{-1}{P_0 + \hat{K}(\mathbf{q})} = \frac{-1}{P_0 + \hat{K}(\mathcal{Q})} + \frac{\hat{K}(\mathbf{q}) - \hat{K}(\mathcal{Q})}{[P_0 + \hat{K}(\mathcal{Q})]^2} - \dots \quad (\text{A26})$$

This form is illuminating since in the case of predominantly ferromagnetic couplings of the  $\sigma$  spins one has, for rapidly decaying interactions in a large system,

$$\hat{K}(\mathbf{q}) = \hat{K}(\mathcal{Q})[1 - R_0^2 q^2 + O(q^4)], \quad (\text{A27})$$

where  $\hat{K}(\mathcal{Q}) > 0$ , while the (real) length  $R_0$  measures the range of the interactions. By substitution in (A26) we see that the couplings of the  $s$  spins are likewise ferromagnetic with a comparable finite range. Finally, note that in the thermodynamic limit one has

$$Q_{ii} = Q(\mathcal{Q}) = \int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{P_0 + \hat{K}(\mathbf{q})}, \quad (\text{A28})$$

which is necessarily positive as required for a sensible weighting factor.

## APPENDIX B Details of the $\epsilon$ -expansion calculation

In this appendix we examine the derivation of the recursion relations (6.26)-(6.27) and (6.40)-(6.42) for the perturbation-theoretic expansion of the Wilson momentum shell renormalization group near  $d=4$  dimensions and introduce the diagrammatic language that facilitates the calculations. We are concerned here with the LGW reduced Hamiltonian (6.4) which we rewrite, as in (6.11), in the form

$$\bar{\mathcal{H}} = -\frac{1}{2} \int_{\mathcal{g}} \sum_{\mu=1}^n (r + eq^2) \sigma_{\mathcal{g}}^{\mu} \sigma_{-\mathcal{g}}^{\mu} - u \int_{\mathcal{g}_1} \int_{\mathcal{g}_2} \int_{\mathcal{g}_3} \sum_{\mu, \nu=1}^n \sigma_{\mathcal{g}_1}^{\mu} \sigma_{\mathcal{g}_2}^{\mu} \sigma_{\mathcal{g}_3}^{\nu} \sigma_{\mathcal{g}_4}^{\nu} \quad (\text{B1})$$

with the wavevectors restricted by

$$g_1 + g_2 + g_3 + g_4 = 0, \quad (B2)$$

so that 'umklapp' processes are neglected. Likewise, we suppose that the sixth and higher order terms may be neglected initially. The validity of both these approximations, to leading order in  $\epsilon$ , may be checked by computing their effects by the same techniques. However, it is important to note that even if such terms are rigorously absent in the initial, physical Hamiltonian they may be generated and normally will be generated, in the process of successive renormalization. The coefficient  $e$  of the  $q^2$  term will be constrained throughout (by spin rescaling) to be equal to unity. Thus  $\mathcal{H}$  is, effectively, determined only by the two parameters  $r$  and  $u$ .

As discussed in the main text, the first step is to split up  $\mathcal{H}$  as follows

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}^< + \bar{\mathcal{H}}^>_2 - u \bar{\mathcal{H}}^>_4. \quad (B3)$$

Then we must compute the renormalized Hamiltonian,  $\bar{\mathcal{H}}'$ , which is given to second order in  $u$  by

$$\begin{aligned} \bar{\mathcal{H}}' &\approx \{ \bar{\mathcal{H}}^< + \ln(\text{Tr} \langle e^{\bar{\mathcal{H}}^>_2} \rangle) - u \langle \bar{\mathcal{H}}^>_4 \rangle \} \\ &+ \frac{1}{2} u^2 [ \langle (\bar{\mathcal{H}}^>_4)^2 \rangle - \langle \bar{\mathcal{H}}^>_4 \rangle^2 ] + \dots \} \delta_{\mathbf{q}} \Rightarrow \delta'_{\mathbf{q}}. \end{aligned} \quad (B4)$$

Let us start by examining the lower order terms. First, note that  $\bar{\mathcal{H}}^<$  is of the same basic form as the original LGW Hamiltonian except that the momentum integrals are limited to the inner region,  $<$ , of momentum space which contains only  $N' = N/b^d$  spins. However, the spatial rescaling restores the original domain of integration through the transformations

$$\mathbf{q} \Rightarrow \mathbf{q}'/b: \int_{\mathbf{q}}^< \equiv \int_{\mathbf{q}'}^< \frac{d^d \mathbf{q}}{(2\pi)^d} = \int \frac{d^d(\mathbf{q}'/b)}{(2\pi)^d} \equiv b^{-d} \int_{\mathbf{q}'}, \quad (B5)$$

where uninflected momentum integrals run over the full zone. Recalling the spin rescaling,  $\hat{\sigma}_{\mathbf{q}} \Rightarrow \hat{c} \hat{\sigma}'_{\mathbf{q}'}$ , we see that the quadratic part of  $\bar{\mathcal{H}}^<$  transforms as

$$\begin{aligned} -\frac{1}{2} \int_{\mathbf{q}}^< (r + e q^2) \hat{\sigma}_{\mathbf{q}} \cdot \hat{\sigma}_{-\mathbf{q}} &= -\frac{1}{2} \int_{\mathbf{q}'} b^{-d} (r + e q'^2 b^{-2}) \hat{c}^2 \hat{\sigma}'_{\mathbf{q}'} \cdot \hat{\sigma}'_{-\mathbf{q}'}, \\ &= -\frac{1}{2} \int_{\mathbf{q}'} [ (\hat{c}^2 b^{-d} r) + (\hat{c}^2 b^{-d-2} e) q'^2 ] \hat{\sigma}'_{\mathbf{q}'} \cdot \hat{\sigma}'_{-\mathbf{q}'}, \end{aligned} \quad (B6)$$

so that, ignoring possible contributions from higher order terms, etc., the parameters  $r$  and  $e$  are simply renormalized by factors  $\hat{c}^2 b^{-d}$  and  $\hat{c}^2 b^{-d-2}$ , respectively. Notice that the momenta,  $\mathbf{q}'$ , and spins,  $\hat{\sigma}'_{\mathbf{q}'}$ , in (B6) are really

'dummy variables', in the sense that they are to be integrated over, so we may actually drop all the primes in the final expression. Evidently the quartic term similarly generates a leading renormalization factor  $c^4 b^{-3d}$  for  $u$ .

In the second term in (B4) the term  $\exp(\bar{\mathcal{H}}_2^>)$  depends only on spins  $\sigma^>$ , in the outer zone and so when the operation  $\text{Tr}^>$  is applied to it, the result is just a constant contribution which can be ignored for our present purposes since it cannot further affect other interactions. (Nevertheless, if we were concerned to calculate the free energy itself, we would have to retain this constant term in  $\bar{\mathcal{H}}$ .) This completes the calculation of the recursion relations to zeroth order in  $u$ : see (6.26) to (6.28).

The evaluation of the first order term,  $u \langle \bar{\mathcal{H}}_4^> \rangle$ , requires closer attention. First note that in the expression for  $\bar{\mathcal{H}}_4^>$ , which for finite  $N$  is just the multiple sum,

$$\bar{\mathcal{H}}_4^> = \left(\frac{a^{-d}}{N}\right)^3 \sum_{q_1} \sum_{q_2} \sum_{q_3}^> \sum_{\mu, \nu=1}^n \sigma_{q_1}^\mu \sigma_{q_2}^\mu \sigma_{q_3}^\nu \sigma_{q_4}^\nu, \tag{B7}$$

at least one of the four momentum labels  $q_1, q_2, q_3$  or  $q_4 = -q_1 -q_2 -q_3$  must lie in the outer region,  $>$ . We must then evaluate the momentum shell average

$$\langle \bar{\mathcal{H}}_4^> \rangle = \frac{\text{Tr}^> \{ \bar{\mathcal{H}}_4^> e^{\bar{\mathcal{H}}_2^>} \}}{\text{Tr}^> \{ e^{\bar{\mathcal{H}}_2^>} \}}, \tag{B8}$$

where  $\bar{\mathcal{H}}_2^>$  can similarly be written as a sum, namely,

$$\bar{\mathcal{H}}_2^> = -\frac{1}{2} \frac{a^{-d}}{N} \sum_{q}^> \sum_{\mu=1}^n (r + eq^2) \sigma_q^\mu \sigma_{-q}^\mu, \tag{B9}$$

the momenta being all restricted to the outer zone. The trace operation has the explicit form

$$\text{Tr}^> = \prod_{q}^> \prod_{\mu} \int_{-\infty}^{\infty} d\sigma_q^\mu, \tag{B10}$$

which signifies a multiple integral over all the spin components  $\sigma_q^\mu$  with momenta in the outer region,  $>$ . In the thermodynamic limit,  $N \rightarrow \infty$ , this operation becomes a functional integral. However, we may avoid this concept and the question of its proper definition by keeping  $N$  finite and doing the  $n(N-N')$  integrals over all the  $\sigma^>$  components before taking the thermodynamic limit. (Note that in the spatial rescaling steps in (B5) and (B6) it was advantageous to take the thermodynamic limit, as we did implicitly, at an early stage.)

Each term in (B7) can be processed separately through (B8) and the results then added together. Consider the typical term  $\sigma_{q_1}^\mu \sigma_{q_2}^\mu \sigma_{q_3}^\nu \sigma_{q_4}^\nu$ . If all four of the  $q_j$  lie

in the  $\rangle$  zone then the result will be a constant and hence is of no further direct interest to us here (although it contributes to the constant term in  $\bar{\mathcal{H}}'$ ). If either one or three of the  $q_j$  lie in the  $\rangle$  zone the result must vanish. To see this note that the integrals involved in (B8) have the symmetric Gaussian form shown in (6.18) and so vanish by symmetry if an odd power of  $\sigma_q^\mu$  is contributed by  $\bar{\mathcal{H}}_4^\rangle$ . (See further below.) At least one of the  $\sigma^\rangle$  integrals must be odd in the case posed. Thus the only terms that need be considered are those where two of the  $q_j$  are equal in magnitude but opposite in sign and both lie in the  $\rangle$  zone, with the associated spin component indices being the same. This generates a term  $\sigma_{q-g}^\mu \sigma_{-g}^\mu = [\text{Re}(\sigma_q^\mu)]^2 + [\text{Im}(\sigma_q^\mu)]^2$ . The other two  $q_j$ , then belong to the inner zone and will, perforce, also be equal in magnitude and opposite in sign; their associated spin component indices,  $\nu$ , will likewise match. Taking into account all possible combinations that satisfy these criteria, leads to

$$\begin{aligned}
 \langle \bar{\mathcal{H}}_4^\rangle \rangle &= \left(\frac{a^{-d}}{N}\right)^3 \sum_{q_{\langle}} \sum_{q_{\rangle}} \sum_{\mu, \nu=1}^n \{ \langle \sigma_{q_{\langle}}^\mu \sigma_{-q_{\langle}}^\mu \sigma_{q_{\rangle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle \\
 &+ \langle \sigma_{q_{\rangle}}^\mu \sigma_{-q_{\rangle}}^\mu \sigma_{q_{\langle}}^\nu \sigma_{-q_{\langle}}^\nu \rangle \rangle + \langle \sigma_{q_{\langle}}^\mu \sigma_{q_{\rangle}}^\mu \sigma_{-q_{\langle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle \delta_{\mu\nu} \\
 &+ \langle \sigma_{q_{\langle}}^\mu \sigma_{q_{\rangle}}^\mu \sigma_{-q_{\rangle}}^\nu \sigma_{-q_{\langle}}^\nu \rangle \rangle \delta_{\mu\nu} + \langle \sigma_{q_{\rangle}}^\mu \sigma_{q_{\langle}}^\mu \sigma_{-q_{\langle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle \delta_{\mu\nu} \\
 &+ \langle \sigma_{q_{\rangle}}^\mu \sigma_{q_{\langle}}^\mu \sigma_{-q_{\rangle}}^\nu \sigma_{-q_{\langle}}^\nu \rangle \rangle \delta_{\mu\nu} \}. \tag{B11}
 \end{aligned}$$

Now all the spin variables commute. Further, the spins  $\sigma_{\langle}$  are not affected by the  $\text{Tr}^\rangle$  operation and can thus be removed from under the angular brackets. In addition, because of the equivalence of the different components for each spin we have

$$\sum_{\nu=1}^n \langle \sigma_{q_{\rangle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle = n \langle \sigma_{q_{\rangle}}^j \sigma_{-q_{\rangle}}^j \rangle \rangle, \tag{B12}$$

and hence find that

$$\langle \bar{\mathcal{H}}_4^\rangle \rangle = \left(\frac{a^{-d}}{N}\right)^3 \left[ (2n+4) \sum_{q_{\rangle}} \langle \sigma_{q_{\rangle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle \right] \sum_{q_{\langle}} \sum_{\mu=1}^n \sigma_{q_{\langle}}^\mu \sigma_{-q_{\langle}}^\mu, \tag{B13}$$

where we have dropped the  $\langle$  subscripts on  $q$  since they are no longer essential. Note that the combinatorial factor  $(2n+4)$  is of central importance to the final answers!

The next step is to calculate  $\langle \sigma_{q_{\rangle}}^\nu \sigma_{-q_{\rangle}}^\nu \rangle \rangle$ . It is equal to a product of integrals over all the  $\sigma^\rangle$ , divided by a similar product. Cancellation occurs for all these integrals except for those over  $\sigma_{q_{\rangle}}^\nu$  and  $\sigma_{-q_{\rangle}}^\nu$ ; hence the result is

$$\langle \sigma_{\mathbf{q}}^{\nu} \sigma_{-\mathbf{q}}^{\nu} \rangle = \frac{\iint d\sigma_{\mathbf{q}}^{\nu} d\sigma_{-\mathbf{q}}^{\nu} \sigma_{\mathbf{q}}^{\nu} \sigma_{-\mathbf{q}}^{\nu} \exp[-\frac{1}{2}(\frac{a}{N})^{-d}(r+eq^2)(\sigma_{\mathbf{q}}^{\nu} \sigma_{-\mathbf{q}}^{\nu} + \sigma_{-\mathbf{q}}^{\nu} \sigma_{\mathbf{q}}^{\nu})]}{\iint d\sigma_{\mathbf{q}}^{\nu} d\sigma_{-\mathbf{q}}^{\nu} \exp[-\frac{1}{2}(\frac{a}{N})^{-d}(r+eq^2)(\sigma_{\mathbf{q}}^{\nu} \sigma_{-\mathbf{q}}^{\nu} + \sigma_{-\mathbf{q}}^{\nu} \sigma_{\mathbf{q}}^{\nu})]} \quad (\text{B14})$$

The integration over pairs of complex conjugate spins  $\sigma_{\mathbf{q}}^{\nu}$  and  $\sigma_{-\mathbf{q}}^{\nu} = (\sigma_{\mathbf{q}}^{\nu})^*$  is carried out by making use of the equivalence

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\sigma_{\mathbf{q}}^{\nu} d\sigma_{-\mathbf{q}}^{\nu} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d(\text{Re}\sigma_{\mathbf{q}}^{\nu}) d(\text{Im}\sigma_{\mathbf{q}}^{\nu}). \quad (\text{B15})$$

By this means we arrive at a product of two separate Gaussian integrals over real variables divided by two other similar integrals: all are readily evaluated and yield

$$\langle \sigma_{\mathbf{q}}^{\nu} \sigma_{-\mathbf{q}}^{\nu} \rangle = \frac{Na^d}{(r+eq^2)}. \quad (\text{B16})$$

Substituting back into (B13), letting  $N \rightarrow \infty$  and replacing the momentum sums by integrals yields

$$\langle \bar{\mathcal{H}}_4 \rangle = 2(n+2) \int_{\mathbf{q}}^{\rangle} \frac{1}{(r+eq^2)} \int_{\mathbf{q}}^{\langle} \vec{\sigma}_{\mathbf{q}} \cdot \vec{\sigma}_{\mathbf{q}}. \quad (\text{B17})$$

Finally, momentum and spin rescaling introduce a factor  $\hat{c}^2 b^{-d}$  as discussed earlier. Bearing in mind the factor  $-u$  associated with  $\langle \bar{\mathcal{H}}_4 \rangle$  in (B4) and the factor  $-1/2$  in the definition of the quadratic part of  $\bar{\mathcal{H}}$ , we see that to first order in  $u$ , the renormalized Hamiltonian has modified values,  $r'$  and  $e'$ , of  $r$  and  $e$  given by the recursion relations

$$r' = \hat{c}^2 b^{-d} [r + 4u(n+2) \int_{\mathbf{q}}^{\rangle} \frac{1}{(r+eq^2)} + O(u^2)], \quad (\text{B18})$$

$$e' = \hat{c}^2 b^{-d-2} [1 + 0 + O(u^2)], \quad (\text{B19})$$

just as stated in (6.40) and (6.41).

The rest of the calculation proceeds in a similar fashion, with intermediate algebraic expressions of the type displayed in (B11) becoming considerably more complex. However, the combinatorial problem of deciding just which terms can contribute is greatly simplified by the use of graphical or diagrammatic notation. We define, first, the free inverse "propagator" by

$$[G_0^{\mu\nu}(\mathbf{q})]^{-1} = \delta_{\mu\nu}(r+eq^2) \approx \frac{\mu}{\mathbf{q}}. \quad (\text{B20})$$

This carries both a momentum and a spin component index, and serves to represent the

quadratic or "free field" part of the total Hamiltonian. In field-theoretic language  $r$  represents the "bare mass". To represent the quartic part of the Hamiltonian we introduce the

$$\text{four-point vertex} \approx \begin{matrix} q_1^\mu \\ \swarrow \\ \dots \\ \searrow \\ q_2^\mu \end{matrix} \dots \dots \dots \begin{matrix} q_3^\nu \\ \swarrow \\ \dots \\ \searrow \\ q_4^\nu \end{matrix} \quad (B21)$$

This has four incoming lines which carry momentum and spin component indices  $\mu, \mu, \nu$  and  $\nu$  corresponding to the term  $\sigma_1^\mu \sigma_2^\mu \sigma_3^\nu \sigma_4^\nu$  in (B1). Momentum is conserved 'through' a vertex in accord with the condition (B2). Each vertex also carries a "coupling constant" factor  $u$  ( $\equiv u_4$ ).

The process of calculating the renormalized (or "dressed") propagator can now be represented graphically as

$$\left( \overset{\mu}{\underset{g}{\rightarrow}} \overset{\mu}{\leftarrow} \right)' = \left( \overset{\mu}{\underset{g}{\rightarrow}} \overset{\mu}{\leftarrow} \right) + u \left[ \left( \overset{g^\mu}{\underset{-g_\mu}{\rightarrow}} \dots \dots \dots \overset{g^\nu}{\leftarrow} \right) + \left( \left( \circ \right) \dots \dots \dots \left( \leftarrow \right) \right) \right] \quad (B22)$$

$$+ \left[ \left( \overset{g^\mu}{\rightarrow} \dots \dots \dots \overset{-g^\mu}{\leftarrow} \right) + \left( \overset{g^\mu}{\rightarrow} \dots \dots \dots \left( \leftarrow \right) \right) + \left( \left( \rightarrow \right) \dots \dots \dots \left( \leftarrow \right) \right) + \left( \left( \rightarrow \right) \dots \dots \dots \left( \leftarrow \right) \right) \right] + O(u^2).$$

The first order diagrams, here, have been arranged in the same order as the corresponding terms in (B11) to facilitate comparison. Evidently they correspond simply to all possible ways of joining up two "legs" of the four-point vertex with matching spin components and momenta in order to leave a propagator-like term. The rules that must be adhered to in constructing the integrals associated with the allowed diagrams are as follows: Each vertex line is accompanied by a factor  $u$ . Internal lines carry a propagator factor  $G_0(q)$  and (in this renormalization group application) imply integration over the outer zone,  $\int$ . If an "internal line" (i.e., one for which  $g$  is to be integrated) forms a closed loop then its spin index is "free" and can be summed over to yield a factor  $n$ . In higher orders of the perturbation theory there is a factor of  $1/m!$  arising from the expansion of the exponential as in (B4). Diagrams that decompose into disconnected parts, i.e., "separated" or "unlinked" diagrams factorize and then cancel when  $\mathcal{K}'$  is computed. (This is an example of the "linked cluster theorem").

The diagrammatic expansion for the renormalized vertex itself is thus found to be

$$\begin{aligned} \begin{matrix} g_1^\mu & g_3^\nu \\ \langle \dots \rangle & \langle \dots \rangle \\ g_2^\mu & g_4^\nu \end{matrix} &= \begin{matrix} g_1^\mu & g_3^\nu \\ \langle \dots \rangle & \langle \dots \rangle \\ g_2^\mu & g_4^\nu \end{matrix} - \frac{u}{2} \left[ \begin{matrix} g_1^\mu & g^\alpha & g_3^\nu \\ \langle \dots \rangle & \langle \dots \rangle & \langle \dots \rangle \\ g_2^\mu & g'^{\alpha'} & g_4^\nu \end{matrix} + \left( \dots \right) \right] + \dots \\ & \qquad \qquad \qquad 8 \text{ (diagrams)} \times n \text{ (free spin components)} \end{aligned}$$

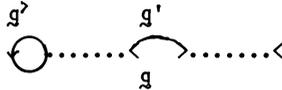
$$+ \left( \begin{matrix} \mu & \nu \\ \langle \dots \rangle & \langle \dots \rangle \\ \mu & \nu \end{matrix} \right) + \dots$$

32 (diagrams) of weight 1

$$+ \left( \begin{matrix} \mu & \nu \\ \langle \dots \rangle & \langle \dots \rangle \\ \mu & \nu \end{matrix} \right) + \dots + O(u^3) \tag{B23}$$

32 (diagrams) of weight 1

Note that diagrams such as



with an "articulation" line or "cut bond" cannot arise,<sup>51</sup> since momentum conservation would require that  $g$  (in  $\langle$ ) =  $-g'$  (in  $\rangle$ ) which is impossible.

The two internal lines in the two-vertex diagrams yield, on integration, a factor

$$\int_{\langle}^{\rangle} \int_{g'}^{\rangle} \frac{1}{(r+eq^2)} \frac{1}{(r+eq'^2)}, \text{ with } g + g' = g_1 + g_2. \tag{B24}$$

A little reflection shows that the renormalized vertex has, in fact, become  $q$ -dependent in that it no longer carries only a constant coupling constant factor,  $u$ , but rather involves a kernel  $u_4(g_1, g_2, g_3)$ . However, we may expand this kernel in powers of the  $g_j$  and associate the coupling constant  $u$  with  $u_4(0,0,0)$ . Likewise then, the renormalized coupling constant  $u'$  is to be associated with  $u'_4(0,0,0)$ . Consequently we can put  $g_1 = g_2 = 0$  so that  $g = -g'$  and the factor for the internal lines thus becomes simply  $\int_{\langle}^{\rangle} \frac{1}{(r+eq^2)^2}$ . After allowing for spin and spatial rescaling, the recursion relation for  $u$  that follows is seen to be

$$u' = c^4 b^{-3d} u \left[ 1 - \frac{1}{2} u (8n + 64) \int_{\langle}^{\rangle} \frac{1}{(r+eq^2)^2} + O(u^2) \right], \tag{B25}$$

in agreement with (6.42). Note that the combinatorial factor  $(8n+24) = 8(n+8)$  directly represents the breakdown of the diagrams in (B24). With practice one learns how to write down such combinatorial factors by inspection for such simple diagrams as here, and by fairly rapid analysis for more complex diagrams like those

that enter in calculation of the  $\epsilon^2$  and  $\epsilon^3$  terms. Reference to the  $\epsilon$ -expansions (6.64), etc. shows that it is just these combinatorial factors that determine the  $n$  and  $\epsilon$  dependence of the various exponents!

The discovery of a  $q$ -dependence in the renormalized vertex is typical of how new terms are generated on renormalization. It indicates that, in principle, such  $q^2, q^4, \dots \sigma^4$ -terms should have been included in the original Hamiltonian, along with  $q^0, q^2, q^4, \dots \sigma^6$  terms and so on. However, on renormalization each  $q^2$  factor would gain an extra renormalization factor  $b^{-2}$ , etc., so that one sees that such terms represent, at least near  $d=4$ , successively more irrelevant critical operators. Nonetheless, it is clear that care and thought are required: blind calculation may lead to a correct answer but an awareness of the general structure of the renormalization group process is a necessary guide if pitfalls are to be avoided!

We saw in (B19) that there is no first order contribution to the renormalization of the coefficient  $e$  which determines the decay exponent  $\eta$ . It is worthwhile recording that the required leading correction comes from the second order propagator diagrams

$$\left( \begin{array}{c} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \\ \text{---} \end{array} \right) + \left( \begin{array}{c} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \\ \text{---} \end{array} \right),$$

$$16\eta \quad + \quad 32$$

which yield the recursion relation (with  $e \equiv 1$ )

$$e' = b^{-\eta} [1 + 16 (n+2) I u^2 \ln b + O(u^3)], \quad (\text{B26})$$

where

$$I = \lim_{b \rightarrow \infty} \frac{128\pi^4}{1 \ln b} \int_{\underline{q}}^{\underline{q}'} \int_{\underline{q}'}^{\underline{q}''} \frac{1}{q^2 q'^2 (q^2 + q'^2)} = \frac{1}{2}. \quad (\text{B27})$$

This recursion relation then yields  $\eta$  correct to order  $\epsilon^2$  as quoted in (6.71). Thus one need not go to third order in  $u$  to find  $\eta$  to  $O(\epsilon^2)$  although this is necessary for the other exponents.

#### APPENDIX C Dimensionality as a Continuous Variable

In the  $(\epsilon=4-d)$ -expansion for critical exponents the spatial dimensionality,  $d$ , is treated as a continuously variable parameter. One way of giving definite meaning to this procedure is based on the observation that the only place that dimensionality enters into the calculations is in performing various integrals which

are of the form  $\int d^d q f(q)$  where, in the simplest case, the integrand,  $f(q)$ , is actually spherically symmetric and therefore a function only of  $q^2$  rather than of the individual components of  $q$ . When  $d$  has a standard integral value the simplest way of doing this type of integral is to transform to hyperspherical coordinates. The integrand depends only on the radial component, and so the integration can be performed immediately over the angular coordinates. Thus, as mentioned in the text of Sec. 6.5, one obtains

$$\int d^d q f(q^2) = C_d \int_0^\infty f(q^2) q^{d-1} dq, \quad (C1)$$

where the area of a unit  $d$ -sphere is given, as in the text, by

$$C_d = 2\pi^{d/2} / \Gamma(\frac{1}{2}d). \quad (C2)$$

The gamma function,  $\Gamma(\frac{1}{2}d)$ , is a well-defined analytic function of its argument so that (C1) is meaningful mathematically even when  $d$  is nonintegral. Thus the extension to arbitrary (even complex!) values of  $d$  is straightforward for functions which are spherically symmetric.

At the next stage one encounters integrals which also involve scalar products such as  $g \cdot p$ , where  $p$  is some reference momentum. Such integrals can be dealt with by the formula

$$\int d^d q f(q^2, g \cdot p) = C_{d-1} \int_0^\infty q^{d-1} dq \int_0^\pi (\sin\theta)^{d-2} d\theta f(q^2, pq\cos\theta). \quad (C3)$$

More generally, following Wilson<sup>52</sup> one only needs the following properties of general  $d$ -dimensional integrals:

(a) Linearity: 
$$\int d^d q [f_1(q) + f_2(q)] = \int d^d q f_1(q) + \int d^d q f_2(q), \quad (C4)$$

(b) Translation Invariance: 
$$\int d^d q f(q + p) = \int d^d q f(q), \quad (C5)$$

(c) Scaling: 
$$\int d^d q f(bq) = b^{-d} \int d^d q f(q), \quad (C6)$$

(d) Normalization: 
$$\int d^d q e^{-q^2} = \pi^{d/2}. \quad (C7)$$

Then, by way of illustration, if one needs an integral such as

$$I_1(p_1, p_2) = \int d^d q \frac{(q \cdot p_1)^2 (q \cdot p_2)^2}{r + q^2} \quad (C8)$$

one first uses the identity

$$\frac{1}{r + q^2} = \int_0^\infty e^{-(r + q^2)s} ds, \quad (C9)$$

to reduce the problem to the Gaussian-type integral

$$I_2(p_1, p_2; s) = \int d^d q (g \cdot p_1)^2 (g \cdot p_2)^2 e^{-sq^2}. \quad (C10)$$

But this can be obtained by differentiating the generating function

$$I_0(\alpha; s) = \int d^d q e^{-sq^2 + g \cdot \sum_i \alpha_i p_i}, \quad (C11)$$

with respect to  $\alpha_1$  and  $\alpha_2$  twice and setting all the  $\alpha_i$  to zero. On the other hand the generating function may be evaluated for general  $d$ , using (C4) to (C7), simply as:

$$I_0(\alpha; s) = s^{-d/2} \pi^{d/2} \exp\left[\left(\sum_i \alpha_i p_i\right)^2 / 4s\right]. \quad (C12)$$

These considerations suffice for field-theoretic applications and hence for the formal developments of  $\epsilon$ -expansions. One may, however, be concerned about the use of a lattice cut off such as enters in, say, the exact solution of the spherical model. If one has nearest neighbor couplings on a hypercubic lattice one then encounters  $d$ -fold integrals like

$$I^{(d)}(z) = \int_{-\pi}^{\pi} \frac{d\theta_1}{2\pi} \dots \int_{-\pi}^{\pi} \frac{d\theta_d}{2\pi} [z + \sum_{j=1}^d \cos \theta_j]^{-1}. \quad (C13)$$

However, by using (C9) and the integral expression for the Bessel function  $J_0(x)$  this can be transformed to

$$I^{(d)}(z) = \int_0^\infty [J_0(s)]^d e^{-zs} ds, \quad (C14)$$

which is again well-defined for general  $d$ . The  $d$ -dependent critical exponents obtained for the spherical model this way agree precisely (to the orders of  $\epsilon$  available) with the  $\epsilon$  expansion expressions evaluated with  $n \rightarrow \infty$ . (see sec. 4.6).

One may discuss the continuation of dimensionality for lattice models more generally. It is natural to restrict attention to hypercubic lattices which a moment's thought shows have a coordination number  $2d$ . This statement, of course, immediately extends to nonintegral values of  $d$ !. To see how to proceed further, consider, to be concrete, the susceptibility of a spin 1/2 Ising model with nearest neighbor coupling of strength  $J$ . The susceptibility may be expressed in terms of the spin-spin correlation functions  $\langle s_{\underline{0}} s_{\underline{R}} \rangle$  between sites  $\underline{0}$  and  $\underline{R}$  as

$$\bar{\chi}(T) = \sum_{\underline{R}} \langle s_{\underline{0}} s_{\underline{R}} \rangle, \quad (C15)$$

where, as usual, the angular brackets denote the statistical expectation defined here by

$$\langle A \rangle = 2^{-N} \sum_{\{s_i = \pm 1\}} (A(s)) \prod_{(ij)} e^{K s_i s_j} / 2^{-N} \sum_{\{s_i = \pm 1\}} \prod_{(ij)} e^{K s_i s_j}, \quad (C16)$$

the sums running over all spin states of the lattice. In the high temperature limit,  $K = J/k_B T \rightarrow 0$ , one can, as pointed out in Sec. 4.5, expand the exponential factors in powers of  $K$  as

$$e^{K s_i s_j} = 1 + K s_i s_j + \frac{1}{2!} K^2 s_i^2 s_j^2 + \dots \quad (C17)$$

Each power of  $K$  in the full expansion of (C17) is clearly associated with a nearest neighbor lattice bond. When calculating the susceptibility an extra pair of spins,  $S_0$  and  $S_R$ , will appear in the expression for the numerator of (C16). The resulting expansion in terms of multiple spin products must be summed over all possible spin configurations. The contribution of any given product of spins may then be evaluated by using the identity

$$\begin{aligned} \frac{1}{2} \sum_{s=\pm 1} s^k &= 1, \text{ for } k \text{ even,} \\ &= 0, \text{ for } k \text{ odd.} \end{aligned} \quad (C18)$$

Finally, by collecting up similar terms one sees that the expansion for the susceptibility can be written for any lattice in the diagrammatic form

$$\begin{aligned} \bar{\chi} &= 1 + a_1 [\text{---}] K + a_2 [\text{---}] K^2 + a_3 K^3 [\text{---}] \\ &+ (a_{4,1} [\text{---}] + a_{4,2} [\text{---}]) K^4 + \dots \quad (C19) \end{aligned}$$

The coefficients  $a_1 = 1$ ,  $a_2$ ,  $a_3$ , etc. depend on the topology of the associated diagram, representing bonds on the lattice, but are independent of the lattice structure (or dimensionality) which, in turn, is embodied only in the values ascribed to the graph embedding constants,  $[\text{---}]$ ,  $[\text{---}]$ , etc. (The reader should go through the derivation of the first few terms to see how this works (See also Sec. 4.5).

Now, more explicitly,  $[\text{---}]$  denotes the number of bonds per lattice site. In a  $d$ -dimensional hypercubic lattice, this is evidently  $[\text{---}] = \frac{1}{2} (2d) = d$ , bearing in mind that each nearest-neighbor bond is shared between two lattice sites. In a similar way,  $[\text{---}]$  denotes the number of chains of length two bonds (per lattice site) where successive bonds of the chain must not lie on top of one another: since there are  $\frac{1}{2} (2d)$  choices per site for the first bond and  $(2d-1)$  remaining choices for placing the second bond at one end of the first bond, we obtain  $[\text{---}] = d(2d-1)$ .

Likewise, the square yields  $[\square] = 2d(d-1)$ , and so on. Evaluation at  $d=3$  then yields the series quoted in (4.19).

For completeness we digress a moment to recall that the Ising series presented in (4.19) is given in terms of the variable

$$v = \tanh K = K - \frac{1}{3}K^3 + \dots, \quad (C20)$$

rather than in powers of  $K$  directly. The usefulness of this variable, in fact, arises directly from the diagrammatic or graphical expansion technique: thus, for Ising variables, for which  $s_i s_j$  can take only the two values  $+1$  or  $-1$ , it is simpler to replace the infinite expansion (C17) by the two-term identity

$$e^{Ks_i s_j} = (\cosh K) [1 + vs_i s_j], \quad (C21)$$

which is easily checked. In making expansions of (C16) in powers of  $v$ , each bond now appears only once, with weight  $v$ , rather than multiply with weights  $K, K^2, K^3, \dots$  as entailed in the use of (C17).

To return to the general theme, it should now be clear that even the most complicated diagram entering in a graphical expansion will have embedding constants or weights that are just polynomials in the dimensionality  $d$ . It follows that each term in the high temperature expansion of  $\bar{\chi}(T)$  can be analytically continued to arbitrary values of  $d$ . Thus, at least while the series converges, the susceptibility itself can also be defined for continuous dimensionality.

The same procedure works for all other properties. This lattice definition and the prescription of introducing continuous  $d$  through various integrals do not obviously agree in general (and no such proofs have been presented). Wherever they have been tested, however, the different prescriptions appear to coincide and, in particular, it is reasonable to expect that they will all yield the same results in the critical region.

#### APPENDIX D Hyperscaling and Dangerous Irrelevant Variables

Consider the hyperscaling relation  $d\nu = 2-\alpha$ . This relation was obtained in Sec. 5.5.4 from the renormalization of the correlation length according to

$$\xi[\bar{\mathcal{K}}] = b\xi[\bar{\mathcal{K}}'], \text{ with } \xi \sim t^{-\nu}, \quad (D1)$$

which merely represents the basic rescaling of lengths, and of the free energy according to

$$f[\bar{\mathcal{K}}] = b^{-d}f[\bar{\mathcal{K}}'], \text{ with } f_{\text{sing.}} \sim t^{2-\alpha}. \quad (D2)$$

The question is: "How can the arguments go wrong, as the breakdown of hyperscaling in large dimensionalities implies must happen?". There are, in fact, various mechanisms by which hyperscaling can fail. To explain the most likely mechanism, which does not actually violate the basic structure of the renormalization group theory<sup>53</sup> let us recapitulate the argument for hyperscaling.

As seen in Sec. 5.5, near a fixed point the free energy depends on a number of scaling fields,  $g_1, g_2, \dots$ , in terms of which it should scale asymptotically in the form

$$f(g_1, g_2, \dots) \approx b^{-d} f(b^{\lambda_1} g_1, b^{\lambda_2} g_2, \dots). \quad (D3)$$

If we make the standard choice and identification

$$b^{\lambda_1} = 1/g_1 \approx 1/t, \quad (D4)$$

we obtain

$$f(g_1, g_2, \dots) \approx t^{d/\lambda_1} f(1, \frac{g_2}{\lambda_2/\lambda_1}, \frac{g_3}{\lambda_3/\lambda_1}, \dots). \quad (D5)$$

The natural example to consider is provided by the simplest continuous spin ferromagnet where  $g_2 \approx H$  and  $\lambda_2/\lambda_1 = \Delta$ , while  $g_3 = u$  represents the coefficient of the quartic spin term [see Sec. 6.1] with  $\lambda_3/\lambda_1 = \phi_3 \equiv \phi$ . More generally, however, we need not specify the nature of  $u$ . Then we have

$$f(t, H, u) \approx t^{d/\lambda_1} Y_0\left(\frac{H}{t^\Delta}, \frac{u}{t^\phi}\right), \quad \text{with} \quad Y_0(y, z) \approx f(1, y, z, 0, 0, \dots), \quad (D6)$$

where, for simplicity, we now ignore all further variables which we thus assume are "harmless" irrelevant variables.

The scaling exponent  $\phi$  may, in principle, be positive, negative or zero. If it is positive then  $u$  is actually a relevant variable and its flow under renormalization is away from the fixed point selected. One is then dealing with some sort of multicritical situation which is not pertinent to the present issue. On the other hand, if  $\phi$  is negative  $u$  is formally irrelevant and on approach to the critical point, one has

$$\frac{u}{t^\phi} = ut^{|\phi|} \rightarrow 0 \quad \text{as} \quad t \rightarrow 0. \quad (D7)$$

Therefore however large  $u$  was initially, the scaled combination  $u/t^\phi$  becomes arbitrarily small asymptotically close to the critical point, and so, formally, one has

$$f \approx t^{d/\lambda_1} Y_0\left(\frac{H}{t}, 0\right) \approx t^{2-\alpha} Y\left(\frac{H}{t}\right), \quad (D8)$$

as argued in Sec. 5.5.4. On making the identification

$$2 - \alpha = d/\lambda_1, \quad (D9)$$

and using the general result  $\nu = 1/\lambda_1$  [see (5.96)] we arrive at the hyperscaling relation  $d\nu = 2 - \alpha$ . Evidently the asymptotic scaling function is given by

$$Y(y) \approx Y_0(y, 0). \quad (D10)$$

Now this analysis relies implicitly on the assumption that  $Y_0(y, 0) = f(1, y, 0, \dots)$  has a well-defined value. It may happen, however, that the full function  $Y_0(y, z)$  actually diverges when  $z \rightarrow 0$ . Note that the fact that  $u$  is an irrelevant variable in no way excludes this possibility! To examine the likely consequences of such a situation let us postulate a simple power law divergence of the form

$$Y_0(y, z) \approx \frac{W(y)}{z^\mu} \text{ as } z \rightarrow 0^+ \text{ with } \mu > 0. \quad (D11)$$

An irrelevant variable,  $u$ , giving rise to this type of behavior is characterized as a dangerous irrelevant variable. Substituting this assumption into (D6) and letting  $u|\phi| \rightarrow 0$  as  $t \rightarrow 0$  now yields

$$f \approx t^{d/\lambda_1} W\left(\frac{H}{t}\right)/u^\mu t^\mu |\phi| \approx t^{(d/\lambda_1) - \mu} |\phi| \tilde{Y}\left(\frac{H}{t}\right), \quad (D12)$$

where  $\tilde{Y}(y) = W(y)/u^\mu$  evidently represents a new asymptotic scaling function. Interpretation of this new behavior in terms of the standard thermodynamic exponents (still accepting  $\nu = 1/\lambda_1$ ) yields the modified relation

$$2 - \alpha = d\nu - \mu |\phi|. \quad (D13)$$

This clearly represents a breakdown of the original hyperscaling relation! Notice, nevertheless, that the renormalization group framework has been preserved intact: the only flaw in the original argument was a failure to recognize and allow for possible singular behavior of the scaling function.

But how far-fetched is the idea of a scaling function diverging as in (D11)? The answer is "Not at all!". Indeed, when  $u$  represents the coefficient of the  $s^4$  term in a continuous spin model, just such a divergence is found when one calculates the form of the free energy scaling function (for nonzero but small  $u$ ) above four

dimensions. Since the  $us^4$  term with  $u > 0$  is essential for the convergence of the partition function below the mean field critical temperature a divergence as  $u \rightarrow 0$  is hardly very surprising. Nevertheless, the actual behavior is, in fact, a little more subtle even than supposed in (D11). What one finds, first, are the renormalization group eigenvalues

$$\lambda_1 = 2, \lambda_2 = \frac{1}{2}d + 1, \text{ and } \lambda_3 = 4 - d, \quad (\text{D14})$$

which pertain to the Gaussian fixed point studied in Sec. 6.4 which is stable for  $d > 4$  since  $\lambda_3 \equiv \lambda_u$  and hence,

$$\phi = \lambda_3/\lambda_1 = -\frac{1}{2}(d - 4), \quad (\text{D15})$$

are then negative. The standard renormalization group exponent identifications yield  $\nu = 1/\lambda_1 = \frac{1}{2}$ , which is the expected classical value, but also

$$2 - \alpha = \frac{d}{\lambda_1} = \frac{1}{2}d, \text{ and } \Delta = \frac{\lambda_2}{\lambda_1} = \frac{1}{4}d + \frac{1}{2}, \quad (\text{D16})$$

neither of which correspond to the classical values,  $\alpha = 0$  and  $\Delta = 3/2$ . However, the scaling function  $Y_0(y, z)$  entering (D6) does behave in a singular manner when  $z \rightarrow 0$ : specifically one finds

$$Y_0(y, z) = \frac{1}{z} W_0(y z^{\frac{1}{2}}), \quad (\text{D17})$$

where  $W_0(w)$  is a well-behaved function. This resembles the postulate (D11) with  $\mu=1$  and so, via (D13), yields the 'operative' or observed critical exponent

$$2 - \alpha = d\nu - \mu|\phi| = \frac{1}{2}d - \frac{1}{2}(d-4) = 2. \quad (\text{D18})$$

Thus we obtain  $\alpha = 0$  which is now in accord with the classical predictions (and, of course, violates hyperscaling).

Evidently, then,  $u$  is a dangerous irrelevant variable at the Gaussian fixed point when  $d > 4$ . Further, however,  $u$ , in the guise of the scaled variable  $z = ut^{|\phi|}$ , also enters as a factor in the argument of the scaling function  $W_0$ . This argument thus becomes

$$w = y z^{\frac{1}{2}} = \frac{H}{\lambda_2/\lambda_1} \left( u^{\frac{1}{2}} t^{\frac{1}{2}|\phi|} \right) = \frac{Hu^{\frac{1}{2}}}{t^{\Delta}}, \quad (\text{D19})$$

where now the operative scaling exponent for the ordering field is seen to be

$$\Delta = \frac{\lambda_2}{\lambda_1} - \frac{1}{2}|\phi| = \frac{1}{4}d + \frac{1}{2} - \frac{1}{4}(d-4) = \frac{3}{2}. \quad (\text{D20})$$

This contrasts sharply with (D15) but agrees with the classical prediction!

The moral of this story is that the standard scaling relations for critical exponents depend, in their derivation, on assumptions, usually left tacit, about the nonsingular or nonvanishing behavior of various scaling functions and their arguments. In many cases these assumptions are valid and may be confirmed by explicit calculation (or other knowledge) but in certain circumstances they may fail, in which case an exponent relation may change its form. Other nontrivial cases of dangerous irrelevant variables are known so that the phenomenon, although not common, is not truly exceptional.

#### BIBLIOGRAPHY

We list here some reviews and monographs as suggestions for further reading and references concerning renormalization group theory and its applications. Some are at a more introductory and less technical level than the present lectures: these are listed first. Others follow roughly in order of increasing technical sophistication.

1. B. Widom in "Fundamental Problems in Statistical Mechanics" vol. III, Ed. E. V. G. Cohen (North Holland, 1975) pp. 1-45.
2. P. Pfeuty and G. Toulouse, "Introduction to Renormalization Groups and Critical Phenomena" (Wiley, 1977).
3. M. E. Fisher, Rev. Mod. Phys. 46, 597-616 (1974).
4. D. J. Wallace and R. K. P. Zia, Repts. Prog. Phys. 41, 1-85 (1978).
5. A. Z. Patashinskii and V. L. Pokrovskii, "Fluctuation Theory of Phase Transitions" (Pergamon Press, 1979).
6. L. E. Reichl, "A Modern Course in Statistical Physics", (Univ. of Texas Press, 1980), Chap. 4.
7. A. P. Young in "Ordering in Strongly Fluctuating Condensed Matter Systems", Ed. T. Riste (Plenum Press, 1980) pp. 11-31, 271-284.
8. S.-K. Ma, "Modern Theory of Critical Phenomena" (Benjamin, 1976).
9. K. G. Wilson, Rev. Mod. Phys. 47, 773-840 (1975).
10. K. G. Wilson and J. Kogut, Physics Reports 12 C (1974).
11. B. I. Halperin in "Physics of Low-Dimensional Systems", Eds. Y. Nagaoka and S. Hikami (Publ. Office, Prog. Theoret. Phys., Kyoto, 1979).
12. Articles by A. Aharony, by E. Brezin, J. C. Le Guillon and J. Zinn-Justin, by Th. Niemeijer and J. M. J. van Leeuwen, by F. J. Wegner and by others in "Phase Transitions and Critical Phenomena", vol. 6, Eds. M. S. Green and C. Domb (Academic Press, 1976). Other volumes in this series provide a broad coverage of other aspects of critical phenomena.
13. D. J. Amit, "Field Theory, the Renormalization Group and Critical Phenomena" (McGraw-Hill, 1978).

## References

1. See First Book of Kings, Chap. 7, verse 23.
2. For example, see the famous pictures reproduced in H. E. Stanley, "Introduction to Phase Transitions and Critical Phenomena" (Oxford Univ. Press, 1971) Fig.1.7.
3. D. Balzarini and K. Ohrn, Phys. Rev. Lett. 29, 840 (1972); D. Balzarini, Can. J. Phys. 52, 499 (1974).
4. M. J. Sienko, J. Chem. Phys. 53, 566 (1970).
5. M. I. Bagatskii, A. V. Voronel' and V. G. Gusak, Sov. Phys. JETP 16, 517 (1963).
6. G. Ahlers, Phys. Rev. Lett. 24, 1333 (1970).
7. B. J. Lipa, unpublished work at Stanford University (1981).
8. See, for example H. E. Stanley loc. cit. Figs. 11.4 and 11.5 or M. E. Fisher, Rep. Prog. Phys. 30, 615 (1967) Fig. 18, etc.
9. See, for example, H. B. Tarko and M. E. Fisher, Phys. Rev. B 11, 1217 (1975) which lists data for various universal and nonuniversal parameters of two and three-dimensional Ising models.
10. Based on the general scaling treatment presented in Proc. Nobel Symposium 24, Eds. B. Lundqvist and S. Lundqvist (Academic Press, 1974) pp. 16-37.
11. Frenkel's philosophy was expressed more fully in a 1946 review article devoted to the theory of metals. I came across his formulation two decades ago in I. E. Tamm's obituary of Yakov Il'ich Frenkel which appeared in Soviet Physics Uspekhi 5, No. 2, Sept./Oct. (1962).
12. See e.g. the review "Simple Ising models still thrive": M. E. Fisher, Physica 106A, 28 (1981) and references therein.
13. C. Domb, Adv. Phys. 9, 149 (1960).
14. See e.g. Domb loc. cit.; C. Domb and M. F. Sykes, J. Math. Phys. 2, 63 (1961); M. E. Fisher, Rept. Prog. Phys. 30, 615 (1967) Figs. 13, 14; D. S. Gaunt and A. J. Guttmann, in "Phase Transitions and Critical Phenomena", Vol. 3, Eds. C. Domb and M. S. Green (Academic Press, London, 1974).
15. See e.g. J. Zinn-Justin, J. Physique 42, 783 (1981). J.-H. Chen, M. E. Fisher and B. G. Nickel, Phys. Rev. Lett. 48, 630 (1982).
16. See R. Balian and G. Toulouse, Phys. Rev. Lett. 30, 544 (1973); M. E. Fisher, Phys. Rev. Lett. 30, 679 (1973).
17. We follow D. R. Nelson and M. E. Fisher, Ann. Phys. (N.Y.) 91, 226 (1975).
18. Various concrete definitions of the correlation length and their inter-relations are discussed in an Ising model context by M. E. Fisher and R. J. Burford, Phys. Rev. 15b, 583 (1967) and M. B. Tarko and M. E. Fisher, Phys. Rev. B 11, 1217 (1975). In most simple critical phenomena all definitions agree up to a constant factor as the critical point is approached. In more complex situations, however, there may be different correlation lengths for different degrees of freedom even though (5.62) is maintained. Sometimes, as in dipolar interactions in Ising systems, two differently behaving correlation lengths are associated with orthogonal spatial directions in which case one may also need to generalize (5.62) to allow for different rescaling factors, say  $b_{\parallel}$  and  $b_{\perp}$ , in different directions.

19. See R. B. Griffiths, *Physica* 106A, 59 (1981); R. B. Griffiths and P. A. Pearce, *J. Stat. Phys.* 20, 499 (1979).
20. F. J. Wegner, *J. Phys. C* 7, 2098 (1974).
21. L. P. Kadanoff, *Physics* 2, 263 (1966).
22. M. J. Buckingham in *Critical Phenomena*, Eds., M. S. Green and J. V. Sengers (National Bureau of Standards Misc. Publ. 273, Washington, D. C., 1966), p. 95.
23. See also P. Seglar and M. E. Fisher, *J. Phys. C* 13, 6613 (1980).
24. V. L. Ginzburg, *Sov. Phys. Solid St.* 2, 1824 (1960).
25. L. P. Kadanoff *et al.*, *Rev. Mod. Phys.* 39, 395 (1967).
26. P. C. Hohenberg in *Fluctuations in Superconductors* edited by W. S. Goree and F. Chilton (Stanford Research Inst., Menlo Park, 1968) p. 305.
27. J. Als-Nielsen and R. J. Birgeneau, *Amer. J. Phys.* 45, 554 (1977).
28. Discovered originally by K. G. Wilson and M. E. Fisher, *Phys. Rev. Lett.* 28, 240 (1972) where the  $\epsilon$ -expansion was introduced.
29. E. Brezin, J. C. LeGuillon, J. Zinn-Justin and B. G. Nickel, *Phys. Lett.* 44A, 227 (1973).
30. A. A. Vladimirov, D. I. Kazakov, and O. Tarasov, *Sov. Phys. JETP* 50, 521 (1979).
31. M. E. Fisher, S.-K. Ma, and B. G. Nickel, *Phys. Rev. Lett.* 29, 917 (1972).
32. M. E. Fisher and D. R. Nelson, *Phys. Rev. Lett.* 32, 1352 (1974).
33. M. E. Fisher and A. Aharony, *Phys. Rev. Lett.* 30, 559 (1973).
34. A. I. Larkin and D. E. Khmel'nitskii, *Sov. Phys. JETP* 29, 1123 (1969) discovered some of these features before the general development of renormalization group theory.
35. See also A. Aharony, *Phys. Lett.* 44A, 313 (1973).
36. E. K. Riedel and F. J. Wegner, *Phys. Rev. Lett.* 29, 349 (1972).
37. R. Abe, *Prog. Theoret. Phys.* 48, 1414 (1972).
38. A. M. Polyakov, *Phys. Lett.* 59B, 79 (1975); see also E. Brezin and J. Zinn-Justin, *Phys. Rev. Lett.* 36, 691 (1976).
39. J. M. Kosterlitz and D. J. Thouless, *J. Phys.* C6, 1181 (1973).
40. J. M. Kosterlitz, *J. Phys.* C7, 1046 (1974).
41. J. V. José, L. P. Kadanoff, S. Kirkpatrick and D. R. Nelson, *Phys. Rev.* B16, 1217 (1977).
42. B. I. Halperin and D. R. Nelson, *Phys. Rev. Lett.* 41, 121 (1978); *E* 41, 519 (1978).
43. A. P. Young, *Phys. Rev.* B19, 1855 (1979).
44. Th. Niemeijer and J. M. J. van Leeuwen, *Physica (Utrecht)* 71, 17 (1974).
45. A. A. Migdal, *Zh. Eksp.-Teor. Fiz.* 69, 810, 1457 (1975).

46. L. P. Kadanoff, *Ann. Phys. (N.Y.)* 100, 359 (1976).
47. See e.g., B. I. Halperin in "Collective Properties of Physical Systems" Nobel Symp. No. 24., Eds. B. Lundqvist and S. Lundqvist (Academic Press, 1974); P. C. Hohenberg and B. I. Halperin, *Rev. Mod. Phys.* 49, 435 (1977); J. Tobochnik, S. Sarker, and R. Cordery, *Phys. Rev. Lett.* 46, 1417 (1981).
48. D. R. Nelson, *Phys. Rev.* B11, 3504 (1975).
49. J. Rudnick and D. R. Nelson, *Phys. Rev.* 13, 2208 (1976).
50. K. G. Wilson, *Rev. Mod. Phys.* 47, 773 (1975).
51. A review of graph theoretical terminology is given by J. W. Essam and M. E. Fisher, *Rev. Mod. Phys.* 42, 271 (1970).
52. K. G. Wilson, *Phys. Rev.* D7, 2911 (1973).
53. M. E. Fisher, in "Renormalization Group in Critical Phenomena and Quantum Field Theory: Proceedings of a Conference", Edited by J. D. Gunton and M. S. Green (Temple University, 1974) pp.65-68.