THE RENORMALIZATION GROUP AND THE ε EXPANSION

Kenneth G. WILSON

Institute for Advanced Study, Princeton, N.J. 08540, USA
and Laboratory of Nuclear Studies, Cornell University, Ithaca, N.Y. 14850, USA

and

J. KOGUT

Institute for Advanced Study, Princeton, N.J. 08540, USA

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM

PHYSICS REPORTS (Section C of Physics Letters) 12, No. 2 (1974) 75–200. NORTH-HOLLAND PUBLISHING COMPANY

THE RENORMALIZATION GROUP AND THE ϵ EXPANSION*

Kenneth G. WILSON
Institute for Advanced Study, Princeton, N.J. 08540, USA
and Laboratory of Nuclear Studies, Cornell University, Ithaca, N.Y. 14850, USA†

and

J. KOGUT‡
Institute for Advanced Study, Princeton, N.J. 08540, USA

Received 2 July 1973

Abstract:

The modern formulation of the renormalization group is explained for both critical phenomena in classical statistical mechanics and quantum field theory. The expansion in $\epsilon = 4 - d$ is explained [d is the dimension of space (statistical mechanics) or space-time (quantum field theory)]. The emphasis is on principles, not particular applications. Sections 1–8 provide a self-contained introduction at a fairly elementary level to the statistical mechanical theory. No background is required except for some prior experience with diagrams. In particular, a diagrammatic approximation to an exact renormalization group equation is presented in sections 4 and 5; sections 6–8 include the approximate renormalization group recursion formula and the Feynman graph method for calculating exponents. Sections 10–13 go deeper into renormalization group theory (section 9 presents a calculation of anomalous dimensions). The equivalence of quantum field theory and classical statistical mechanics near the critical point is established in section 10; sections 11–13 concern problems common to both subjects. Specific field theoretic references assume some background in quantum field theory. An exact renormalization group equation is presented in section 11; sections 12 and 13 concern fundamental topological questions.

*Based on lectures given at Princeton University by K.G.W.
†Present address of both authors.
‡Work supported by the U.S. Atomic Energy Commission, Contract No. AT(11-1)2220.

Contents:

1. Introduction

The purpose of this paper is to discuss recent work on the renormalization group and its applications to critical phenomena and field theory. These ideas are illustrated using the other recent idea of defining critical phenomena and field theory in a space of dimension $4-\epsilon$ (space-time dimension $4-\epsilon$ for field theory) and expanding in powers of ϵ. The emphasis is on critical phenomena; basic ideas will be stressed rather than special results. The presentation is incomplete; this review is not a substitute for current literature. The first section is general and philosophical. Most of the subsequent sections are more pragmatic, being concerned with specific problems and calculations.

Associated with this section there is a list of recent references on the renormalization group and the ϵ expansion.

For a precise list of topics discussed in this paper, see the contents.

1.1. *The renormalization group and coherence problems in physics*

In this section a philosophical discussion of the renormalization group will be given. (One should reread this introduction after studying the rest of the paper.) Toward the end of the first section we begin the review of critical phenomena.

The renormalization group is a method for dealing with some of the most difficult problems of physics. These problems include relativistic quantum field theory, critical phenomena, the Kondo effect [e.g. 1–7] and others. These problems are all characterized by involving a large number of degrees of freedom, in an essential way.

Most of the problems one deals with in physics involve a very large number of degrees of freedom. For example, a crystal, liquid, or gas in macroscopic quantities involves more than 10^{23} electrons, and each coordinate of each electron is a degree of freedom.

In contrast, most theoretical methods work only when one has only one independent variable, i.e. only one degree of freedom. For example, consider the Schrödinger equation for a wave function $\psi(x, y, z)$ for one electron. It is infinitely easier to calculate ψ if one can separate variables in the Schrödinger equation (e.g., write $\psi = \psi_1(r)\psi_2(\theta)\psi_3(\phi)$ in spherical polar coordinates). It is obviously hopeless to compute a wave function for 10^{23} electrons without extraordinary simplifications, justified or otherwise.

Under normal circumstances the 10^{23} or so degrees of freedom can be reduced enormously. The intensive or extensive character of observables (energy is extensive, density is intensive) allows one to reconstruct the properties of a macroscopic system given only a microscopic sample of it. Thus a liquid of only 1000 atoms, say, would probably have approximately the same energy per unit volume and density as the same liquid (at the same temperature and pressure) with 10^{23} atoms.

How far can one reduce the size of a gas, say, without qualitatively changing its properties? The minimum size one can reach without change is called the correlation length. The correlation length ξ depends on the state of the system. For a gas ξ depends on pressure and temperature. In favorable circumstances ξ is only one or two atomic spacings. When ξ is this small there exist a variety of methods for calculating properties of the system: virial expansions, perturbation expansions, Hartree-Fock methods, etc. These methods involve a variety of approximations; but they have one feature in common. They all assume that the properties of matter in bulk can be related to the properties of small clusters of atoms. They involve further assumptions because even a cluster of only three atoms involves too many degrees of freedom to be soluble without considerable simplification.

In special cases the correlation length is much larger than the atomic spacing. The critical point marking the onset of a phase transition is a prime example. Liquid-gas transitions, ferromagnetic transitions, order-disorder transitions in alloys, etc., all exhibit critical points for special values of the thermodynamic variables. (The liquid-gas critical point occurs for critical values T_c and P_c of temperature and pressure.) Precisely at the critical point, ξ is infinite; near the critical point ξ is large.

There are a class of problems, including critical phenomena, which are characterized by having very many degrees of freedom in a region the size of a correlation length. "Very many" means not just 3 or 4, but hundreds or millions if not an infinite number. Other problems of this type besides critical phenomena are the Kondo problem (a magnetic impurity in a metal), the binding of large molecules, and the entire subject of relativistic quantum field theory. In the case of a quantum field, say $\phi(x)$, the field ϕ at each point x is a separate degree of freedom, so any region of finite size contains an infinite number of degrees of freedom. The correlation length of a quantum field is usually the Compton wavelength of the lowest mass particle. In the case of quantum electrodynamics it is the electron Compton wavelength (10^{-11} cm) rather than the photon Compton wavelength (∞) which acts as the correlation length in practice. One can relatively easily relate quantum electrodynamics in a box of size $>10^{-11}$ cm to quantum electrodynamics in all space. Boxes of size $\ll 10^{-11}$ cm cause gross distortions of the interactions of electrons and photons.

The problems listed above are all noted for their intransigence. The binding of molecules is hardly better understood today than it was in 1932. There has been fitful progress in critical phenomena over the last 70 years. There has been sensational progress in *calculating* quantum electrodynamics, but very little progress in *understanding* it; and strong interactions are neither calculable nor understood. The Kondo problem has only recently been studied, and may be closer to solution [2, 3, 7].

Studies of renormalization in quantum field theory and critical phenomena in statistical mechanics both suggest that the behavior of systems with many degrees of freedom within a correlation length is qualitatively different from those with only a few degrees of freedom in a correlation length. The systems we are interested in are usually defined by means of a Hamiltonian, and one would normally have expected that the behavior of the system is determined mainly by the type of interactions present in the Hamiltonian and the strengths of the corresponding coupling constants. This is certainly the case when ξ is small. However, in the problems discussed here, where many degrees of freedom are behaving cooperatively, it appears that the behavior of the system is determined primarily by the fact that there is cooperative behavior, plus the nature of the degrees of freedom themselves. The interaction Hamiltonian plays only a secondary role. Thus, in critical phenomena there has developed the notion of universality, i.e. that all interaction Hamiltonians show the same critical behavior. The idea of universality originates in the "law of corresponding states", which is the hypothesis that all fluids and gases have the same equation of state apart from a renormalization of length and energy scales. For a comparatively recent reference on this law see Guggenheim [8]. Recently the idea of universality has been formulated more generally to relate critical behavior in different systems with arbitrary interactions. See for example Kadanoff [9]. Universality will be discussed further below and in section 12.

The renormalization group approach has two objectives. The first is the practical one of simplifying the task of solving systems with many degrees of freedom contained within a correlation length. The basic idea is the same as in hydrodynamics. In hydrodynamics one introduces new variables such as the density $\rho(x)$ which represents an average over the original microscopic degrees of freedom. All microscopic fluctuations are eliminated in the hydrodynamic equations;

$\rho(x)$ is supposed to show only macroscopic fluctuations. What this means in effect is that the hydrodynamic degrees of freedom are the values of $\rho(x)$ at macroscopically separated points. Thus, there are far fewer hydrodynamic degrees of freedom per unit volume than original microscopic degrees of freedom per unit volume.

The renormalization group approach is similar to the hydrodynamic approach in that again the original microscopic degrees of freedom are replaced by a smaller set of effective degrees of freedom. This is done in steps; in each step the linear density of degrees of freedom is reduced by a factor 2.

Reducing the "density of degrees of freedom" can be realized in several ways. Suppose the original system was a system of spins with spacing L_0. The new effective degrees of freedom might be spins with spacing $2L_0$. Alternatively the new degrees of freedom might be a magnetization density $M(x)$ depending on a continuous variable x (just as hydrodynamics replaces discrete atoms by a density $\rho(x)$). Then to limit the degrees of freedom one imposes the restriction that $M(x)$ only contains fluctuations with wavelengths greater than $2L_0$. These ideas will be formulated more precisely in later sections of this review. (The factor of 2 is arbitrary. In particular, it is often useful to change the density of degrees of freedom only infinitesimally instead of by a factor 2. See later sections.)

This reduction in the degrees of freedom is carried out repeatedly. To start with, the spacing of degrees of freedom is L_0; after 1 step the spacing is $2L_0$, after 2 steps the spacing is $4L_0$, etc. One proceeds with further steps until the separation of effective degrees of freedom is of order the correlation length ξ. In each step one has to construct an effective interaction for the effective degrees of freedom, just as in hydrodynamics one has to construct the hydrodynamic equations for $\rho(x)$. The simplification of the renormalization group lies in the hope that these effective interactions $\mathcal{H}_1$, $\mathcal{H}_2$, etc. are *local* interactions, i.e. the interactions should couple directly only nearby degrees of freedom. This of course assumes that the initial interaction was local, but this is true for the problems of interest here. Thus we assume that the range of interaction in the initial interaction is of order L_0 even when $\xi \gg L_0$. The hope is that the range of interaction in $\mathcal{H}_1$ is of order $2L_0$, the range of interaction in $\mathcal{H}_2$ is of order $4L_0$, etc. The alternative is that the range of interaction is ξ even for $\mathcal{H}_1$, which would be disastrous for the renormalization group approach. This disaster is avoided in the examples that have been worked out so far, but may be a problem in other cases.

If it is true that the range of interaction in $\mathcal{H}_1$ is only of order $2L_0$, then one can imagine (again by analogy to hydrodynamics) that the coupling constants in $\mathcal{H}_1$ can be determined by studying the behaviour of the system confined to a region of size (of order) $2L_0$; i.e. one does not have to discuss regions of size ξ. Thus one no longer has a huge number of degrees of freedom to worry about. To determine the coupling constants in $\mathcal{H}_2$ requires a larger region of size $4L_0$. However, the idea here is to determine $\mathcal{H}_2$ starting from $\mathcal{H}_1$ whose degrees of freedom are spaced by $2L_0$, instead of the original interaction $\mathcal{H}_0$ with spacing L_0. Then one still has a limited number of degrees of freedom to consider. In the same way one constructs $\mathcal{H}_3$ from $\mathcal{H}_2$, $\mathcal{H}_4$ from $\mathcal{H}_3$, etc., until one obtains the $\mathcal{H}_n$ for which $2^n L_0 \sim \xi$. At this stage one has only a few degrees of freedom per correlation length and the problem can hopefully be solved by other methods. In general, when calculating $\mathcal{H}_l$ one has to consider a region of size $2^l L_0$, but one starts from $\mathcal{H}_{l-1}$ whose degrees of freedom have spacing $2^{l-1} L_0$. Thus for any l one can consider a region containing only a few effective degrees of freedom.

Unfortunately, the general renormalization group techniques do not reduce the problem to 1 degree of freedom. One can easily imagine that one needs to consider 60 or more effective degrees of freedom when determining $\mathcal{H}_l$ from $\mathcal{H}_{l-1}$. Suppose for example that the effective degrees of freedom in $\mathcal{H}_{l-1}$ are discrete spins, and that one has to discuss a three-dimensional

cube of width 4 spins: this cube will contain 64 spins. This is not very practical. One must simplify the calculation further so that $\mathcal{H}_l$ can be computed from $\mathcal{H}_{l-1}$ using only *one* degree of freedom. The practical applications of the renormalization group described in later lectures involve either special circumstances (the case $d = 4 - \epsilon$ with small ϵ) or crude approximations (the "approximate recursion formula") such that only one degree of freedom is needed for the calculation of $\mathcal{H}_l$. (See, however, the supplemental list of references.)

The second aim of the renormalization group approach is to explain how the qualitative features of cooperative behavior arise. In the renormalization group framework, these qualitative features result from the iterative character of the renormalization group. Namely, there is a transformation τ which converts $\mathcal{H}_0$ to $\mathcal{H}_1$, $\mathcal{H}_1$ to $\mathcal{H}_2$, etc. The transformation is the same whether one is constructing $\mathcal{H}_1$ from $\mathcal{H}_0$ or $\mathcal{H}_2$ from $\mathcal{H}_1$; in each case one is thinning the degrees of freedom by a factor 2. The only difference is in the lengthscale (L_0 versus $2L_0$) which is easily transformed away. So one has a transformation τ which is to be applied repeatedly:

$$\tau(\mathcal{H}_0) = \mathcal{H}_1, \qquad \tau(\mathcal{H}_1) = \mathcal{H}_2, \qquad \tau(\mathcal{H}_2) = \mathcal{H}_3 \qquad \text{etc.} \tag{1.1}$$

This transformation is to be iterated n times where $2^n L_0$ is of order ξ. When ξ is large, the number of iterations is large.

When one has a transformation τ which is iterated many times, the simplest result we can obtain is that the sequence $\mathcal{H}_l$ approaches a fixed point of τ, namely an interaction $\mathcal{H}^*$ satisfying

$$\tau(\mathcal{H}^*) = \mathcal{H}^*. \tag{1.2}$$

This is what will happen in the examples discussed later in this review.

A fixed point of a transformation is a property of the transformation τ itself. That is, to find possible fixed point Hamiltonians $\mathcal{H}^*$ one must solve the fixed point equation (1.2). These equations make no reference to the choice of initial Hamiltonian $\mathcal{H}_0$.

The possible types of cooperative behavior, in the renormalization group picture, are determined by the possible fixed points $\mathcal{H}^*$ of τ. Suppose for example that there are three fixed points $\mathcal{H}_A^*$, $\mathcal{H}_B^*$, and $\mathcal{H}_C^*$. Then one would have three possible forms of cooperative behavior. If a particular system has an initial interaction $\mathcal{H}_0$, one has to construct the sequence $\mathcal{H}_1$, $\mathcal{H}_2$, etc. in order to find out which of $\mathcal{H}_A^*$, $\mathcal{H}_B^*$, or $\mathcal{H}_C^*$ gives the limit of the sequence. If $\mathcal{H}_A^*$ is the limit of the sequence, then the cooperative behavior resulting from $\mathcal{H}_0$ will be the cooperative behavior determined by $\mathcal{H}_A^*$. In this example the set of all possible initial interactions $\mathcal{H}_0$ would divide into three subsets (called "domains"), one for each fixed point. Universality would now hold separately for each domain. See section 12 for further discussion.

This is how one derives a form of universality in the renormalization group picture. It is not so bold as previous formulations [9]. Experience with soluble examples of the renormalization group transformation for critical phenomena shows that it generally has a number of fixed points, so one has to define domains of initial Hamiltonians associated with each fixed point, and only within a given domain is the critical behavior independent of the initial interaction.

There is no *a priori* requirement that the sequence $\mathcal{H}_l$ approach a fixed point for $l \to \infty$. In principle the sequence for large l could show limit cycle, ergodic or turbulent behavior; in such cases it would be difficult to do much calculation. See [10] for an illustration of ergodic and turbulent behavior. But even if the sequence $\mathcal{H}_l$ does not approach a fixed point, it is unlikely that $\mathcal{H}_n$ for large n is a smooth function of the parameters in $\mathcal{H}_0$. The trouble is that small changes in the parameters in $\mathcal{H}_0$ tend to be amplified or deamplified by the transformation τ, and when τ is iterated many times these amplification or deamplification factors become very large (one would guess of order $\sqrt{n}$ from random walk arguments). Thus if u_0 is a coupling constant in $\mathcal{H}_0$ one would expect large ranges of u_0 which are deamplified ($\mathcal{H}_n$ depends very little on u_0)

separated by small ranges of u_0 which are amplified ($\mathcal{H}_n$ changes very rapidly with u_0). In the case that all sequences approach fixed points, the ranges of u_0 for which amplification occurs are those values of u_0 for which $\mathcal{H}_0$ is near the boundary between two domains, while deamplification occurs when $\mathcal{H}_0$ is well inside a particular domain.

In summary, the basic ideas of the renormalization group are, first, to generate a sequence of effective local interactions $\mathcal{H}_l$. That is, if the spacing of degrees of freedom in $\mathcal{H}_l$ is a_l then the interactions should have range a_l not ξ; and by choosing $a_{l+1} = 2a_l$ one should be able to construct $\mathcal{H}_{l+1}$ from $\mathcal{H}_l$ considering regions of size about a_{l+1} rather than ξ. Secondly, the existence of a transformation τ which is iterated repeatedly to construct $\mathcal{H}_l$ from $\mathcal{H}_0$ suggests that the nature of $\mathcal{H}_l$ for large l will be largely or wholly determined by τ itself rather than $\mathcal{H}_0$, thereby leading to at least a limited form of universality.

To conclude this introduction we outline briefly the history of the renormalization group method. This discussion is not complete and probably not very accurate. Landau [11] proposed a hydrodynamic approach to critical phenomena in the late 1930's. His specific theory gave the same results as earlier mean field theory, which is experimentally false. However, the renormalization group approach is best seen as a more sophisticated realization of Landau's ideas.

The specific ideas of the renormalization group approach appeared in two papers of the early 1950's, in connection with quantum field theory. A formulation of the renormalization group transformation τ was given by Stueckelberg and Petermann [12]. Gell-Mann and Low [13], in a remarkable paper, discussed the idea of a fixed point of the transformation and some of its implications: they showed that a unique value of the bare charge e_0 in quantum electrodynamics would correspond to all (sufficiently small) values of the renormalized charge e. This is the amplification effect discussed earlier. In the limit of infinite cutoff (i.e. in the limit of infinitesimal spacing of the original degrees of freedom) τ is iterated an infinite number of times. Thus one can infinite amplification, e.g. a continuous range of values of e corresponding to one value of e_0. The recipes for renormalization group calculations were reviewed in Bogoliubov and Shirkov [14].

The early work on the renormalization group had two defects. It had no calculable experimental consequences, so no one had to take it seriously. Secondly, the intuitive ideas were encased in a thick shell of formalism; it has required many years to peel off the shell.

In the 1960's an extraordinary paper by Kadanoff [15] on critical phenomena contained an intuitive discussion of the idea of thinning the degrees of freedom. Kadanoff assumed that one could discuss blocks of spins in a ferromagnet as if they were single effective spins with very simple interactions. Kadanoff showed that this assumption implied a set of "scaling laws" relating critical exponents which had been postulated earlier by Widom and others (see [16] for references). Kadanoff did not have a shred of justification for his assumption; the importance of his work was that it provided a simple but ideal picture of what an effective degree of freedom would be and how it would interact. This picture is surely unrealizable in practice. What one does is try to come as close as one can to this picture. One tries to define effective degrees of freedom which are roughly describable as block spins, and interactions which have a simple form (not necessarily Ising-like), at least approximately. In summary, Kadanoff has defined a much more profitable goal to work towards than the elaborate formalisms of the earlier work.

The differential equations of the renormalization group resurfaced in the Kondo problem in the work of Anderson et al. and others [2–7]. Anderson [4] was the first to derive these equations explicitly using the idea of reducing the number of degrees of freedom. The previous field theoretic work had a much less transparent justification.

Until recently all formulations of the renormalization group involved a transformation τ acting on a very restricted space of interactions. Gell-Mann and Low considered the standard electrodynamic interaction with the charge e_0 as the only free parameter. Kadanoff allowed only

nearest neighbor Ising-type interactions. A heavy price is paid for this restriction: one can only define τ when the solution of the problem is known. For example, Gell-Mann and Low define τ in terms of the exact electron propagator, which is not known until electrodynamics is solved. Thus there is only a limited possibility of using the renormalization group to help in solving the theory. A study of the fixed source model of the nucleon, in simplified form, showed that this problem disappears if one is willing to let $\mathcal{H}_l$ contain all possible interactions, not just one or two [17]. It is then relatively easy to construct examples of the transformation τ without solving the theory (see later lectures). Now one pays a different price: one cannot keep track of all possible interactions at once because there are too many of them. So one has to have only a few dominant interactions if the renormalization group approach is to be practical. There can be many more interactions with small coupling constants, as long as these can be handled by perturbation methods. This was precisely the situation in the fixed source model. It is also true in the examples discussed later in this review, but it is not at all certain to be true in general.

Current work on the renormalization group will be summarized below.

There have been other ideas for dealing with problems such as critical phenomena. We mention specifically the Migdal-Polyakov bootstrap approach to critical phenomena and field theory [18–26] and the Johnson-Baker-Willey formulation of electrodynamics [27–29] because these ideas are sometimes confused with the renormalization group approach. There is no transformation like τ in either the Migdal-Polyakov or Johnson-Baker-Willey theories; these authors make no attempt to thin the degrees of freedom. So their ideas cannot be classified as renormalization group methods. In place of consciously reducing the number of degrees of freedom, these authors substitute a prayer that an infinite sum of graphs can be replaced by a calculable subset. See below for further comments.

Another recent development is the Callan-Symanzik equations [30, 31]. Closely related to the original Gell-Mann–Low formulation of the renormalization group, these equations are proving to be valuable tools for analyzing the short distance behavior of Feynman graphs [e.g. 32]. At present the Callan-Symanzik equations are too formal to be practical outside of perturbation theory. They will remain so unless some intuition can be added the way Kadanoff added an intuitive picture to the renormalization group approach.

1.2. *Current references*

Current literature on the renormalization group will now be classified, briefly. See also the supplemental list at the end of this review. First, papers on critical phenomena will be listed. The approximate renormalization group recursion formula is derived in [33]. Numerical solutions of the recursion formula for the Ising case are in [33]. Grover, Kadanoff, and Wegner solved the Heisenberg case numerically [34]. Grover solved the $X-Y$ model numerically [35]. Baker [36] and Dyson [37] have described "hierarchical" models for which the recursion formula is exact. Golner [38] has proposed a form of the recursion formula which gives a non-zero value of η (η is defined in section 2). Golner [39] has solved the recursion formula for an example with three-fold symmetry. The ϵ expansion about 4 dimensions was developed by Fisher and Wilson [40]. A calculation of the scaling properties of all perturbations about the critical point, within the ϵ expansion, is given by Wegner [41]; Fisher and Pfeuty [42] investigate slightly anisotropic Heisenberg models. The recursion formula has been applied to tricritical points (such as ^{3}He–^{4}He mixtures) by Riedel and Wegner [43]. The methods of ref. [33] have been applied to the problem of phase separation by Langer and Bar-on [44].

The Feynman graph method for calculating critical exponents in powers of ϵ (section 8) was developed in [45]. It was applied to the "excluded volume" problem by De Gennes [46].

Nickel [47] has calculated the exponent γ to order ϵ^3. The ϵ expansion for the equation of state near the critical point has been obtained by Brézin, Wallace and Wilson [48], [49]. The charged and neutral Bose gas is discussed by Ma [50]. Exponents in the presence of long range forces are computed by Suzuki [51, 52] and Fisher, Ma, and Nickel [53].

The Gell-Mann–Low formulation of the renormalization group was used by Larkin and Khmel'nitskii [54] to discuss the logarithmic behavior of critical phenomena in four dimensions ($\epsilon = 0$) and uniaxial ferroelectrics in three dimensions. De Pasquale, Di Castro and Jona-Lasinio [55–57] and Migdal [58] have discussed qualitative implications for critical phenomena of the Gell-Mann–Low formulation. Di Castro [59] has used the Gell-Mann–Low theory to confirm some results of [45]. Wegner [60] has given an extensive discussion of all kinds of perturbations about the critical point with emphasis on corrections to the scaling laws, using the modern formulation (see also section 12). Several questions have been discussed by Hubbard [61]; the problem of liquids is discussed by Hubbard and Schofield [62]. The Gaussian model with long range forces is discussed by Niemeijer and Van Leeuwen [63].

The work of Dyson [37], Larkin and Khmel'nitskii [54] and Di Castro and Jona-Lasinio [55–57] preceded the work described in this report. For a good survey of critical phenomena just prior to the new developments, see the entire volume of ref. [56]. The new work is summarized briefly in [64].

Many of the references listed above report specific applications of the renormalization group approach which are omitted from this review.

There is another (earlier, but incomplete) paper (set of lecture notes) on the renormalization group [65]. They concern the ideas of the modern renormalization group approach with less emphasis on calculation than the present report.

The close mathematical analogy between critical phenomena and quantum field theory has been emphasized by Gribov and Migdal [66, 67] and Polyakov [68–70] in terms of Feynman diagrams; Moore [71] compares the Feynman path integral to the partition function; Suri [72] makes the connection using the transfer matrix formalism of statistical mechanics (see section 10).

For a review of work on the renormalization group in quantum field theory using the Gell-Mann–Low formulation see [73]. For a field theoretic formulation of the approximate recursion formula, see [74]. See likewise Golner [75].

The idea of a noninteger dimension d and an expansion about $d = 4$ is a useful theoretical device without physical significance: the physics is only in integer dimension d. Nonintegral dimensions will be introduced by analytic continuation in renormalization group equations (section 4) or Feynman graphs (section 8). The idea of nonintegral dimensions in statistical mechanics is not new: see [76] for example. More recently, Widom has studied $1 + \epsilon$ dimensions [77]. Noninteger d has been used recently in field theory to regularize Feynman graphs [78–80]; a thorough discussion is given in 't Hooft and Veltman [78]. Quantum field theory for noninteger d is discussed in [81] where the ϵ expansion is used to compute anomalous dimensions (see also section 9).

Another expansion technique in both critical phenomena and field theory is the $1/N$ expansion where N is the number of internal components of a spin (critical phenomena) or the number of internal components of a quantum field. Stanley [82] discovered that the limit $N \to \infty$ for a spin system reduces to the soluble spherical model of Berlin and Kac [83]. This was discussed further by Kac and Thomson [84]. Techniques for expanding exponents in $1/N$ were discovered by Abe [85] and Wilson [81]. They have been exploited by Abe and Hikami [86], [87], Suzuki [51], [52], Ma [50], [88], Fisher et al. [53], and Ferrell and Scalapino [89]. Brézin and Wallace [90] and Suzuki [52] have computed the equation of state using this expansion. The expansion is

probably not very useful for $N = 1$ to 3, the cases of most physical interest. This is because the ϵ expansion (valid for any N) has denominators of the form $N + 8$ (see section 8) suggesting the $1/N$ expansion is valid only for $N > 8$. However it is instructive to study models with large N as models in their own right. The application of the $1/N$ expansion to field theories in less than four space-time dimension is discussed in [81].

There is a special situation where the problem of a large number of degrees of freedom within a correlation length is trivial, namely free field theories (quantum field theory) or the Gaussian model in statistical mechanics (see section 3). These theories are trivial because the interaction can be diagonalized in momentum space. There exist a variety of graphical methods for treating small perturbations about these trivial theories. The ϵ expansion makes it possible to do successful calculations in critical phenomena by perturbation methods. There are now a bewildering variety of methods for setting up the ϵ expansion: besides the methods discussed in this paper, De Castro [59] has used the old Gell-Mann–Low methods and Mack [26] has used the Migdal–Polyakov approach. The new renormalization group method is emphasized in this paper because of its potential for handling nonperturbative problems (see refs. [33, 60] and sections 12 and 13). By comparison, for problems which cannot be reduced to a few diagrams, it is hopeless to calculate anything using the Gell-Mann–Low theory. Even for questions of principle the Gell-Mann–Low theory is not satisfactory because it involves assumptions which are technical and not physically motivated [73]. It is hard to believe these assumptions are always valid, although they appear to be true order by order in a diagrammatic expansion [e.g. 73]. The Migdal–Polyakov bootstrap (with conformal invariance assumed) is more interesting although not very practical unless only a few diagrams are important. For field theorists an interesting and unique feature of the Migdal–Polyakov bootstrap is Polyakov's "stream unitarity" [91].

The next seven sections are concerned with critical phenomena. This is partly because this is the most successful application of the renormalization group approach. However, critical phenomena provides the simplest and clearest example of the problem of many degrees of freedom within a correlation length. The experiments are precise and relatively unambiguous and test the fundamental aspects of the theory. The simplest models like the Ising model have no complications other than the basic problem of the large correlation length. The study of critical phenomena gives one an overview of the problem of many degrees of freedom within a correlation length that cannot be obtained any other way and is essential to understand the field theoretical applications discussed in sections 9–13.

There are excellent reviews of experiment and previous theory of critical phenomena (see section 2); this report will be devoted to explaining the renormalization group approach. Enough background will be supplied so one can read this review without previous background in critical phenomena, but the relation of the renormalization group to experiment and prior theory will not be discussed in detail.

Sections 11, 12, and 13 are of importance to statistical mechanics; but the main interest from section 9 on is quantum field theory. They depend heavily on the previous sections. They are part of a series of papers [92, 73] explaining the renormalization group approach as an alternative to canonical field theory. The ultimate aim is to produce a field theory of strong interactions, but breakthroughs are still required both in fundamental theory and techniques of calculation before the aim can be realized. Meanwhile, the study of the renormalization group approach has provided, as a byproduct, useful ideas which have been incorporated into more phenomenological approaches to strong interactions [92, 93]. Particle physicists who have studied the ideas of this paper find that it takes considerable effort over a period of time to understand them; they provide a stimulating point of view on the problems of quantum field theory; and the outlook for the future of these ideas is at present uncertain. See section 14.

1.3. *Elementary facts about the Ising model*

Let us begin by discussing the elementary features of the Ising model. (For the history of the Ising model, see [94].) The model has served in the past as a description of ferromagnetism but probably also constitutes a model of relativistic field theory. We imagine a cubic lattice of points, $n = (n_1, n_2, n_3)$ in three dimensions, say, and attach to each lattice point a spin variable s_n. We suppose that s_n takes on only the discrete values ± 1, so that it is not treated as a full-fledged 3-dimensional quantum mechanical spin variable. If the lattice consists of N sites, the system clearly possesses 2^N possible states. Each of these spin configurations will have a particular energy. If only nearest neighbor spin-spin interactions are allowed the Hamiltonian is,

$$H = -J \sum_n \sum_{\hat{i}} s_n s_{n+\hat{i}} \tag{1.3}$$

where $\{\hat{i}\}$ are the unit vectors for each axis on the lattice (fig. 1.1). Furthermore, if the lattice is

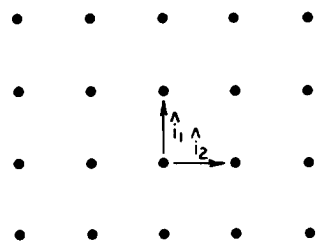

Fig. 1.1. A two-dimensional lattice, with unit vectors for each axis.

emersed in an external magnetic field B, the Hamiltonian acquires an additional term,

$$H = -J \sum_n \sum_{\hat{i}} s_n s_{n+\hat{i}} + \mu B \sum_n s_n \tag{1.4}$$

where μ is a certain gyromagnetic ratio. The thermodynamics of the system can be obtained from the partition function

$$Z = \sum_{\{\text{configurations}\}} e^{-H/kT} \tag{1.5}$$

where the sum runs over all possible spin configurations of the lattice, T is the temperature, and k is Boltzmann's constant. It is also convenient to introduce a free energy F,

$$F = -kT \ln Z. \tag{1.6}$$

Consider also the magnetization per lattice site,

$$M = \frac{1}{N} \sum_n \frac{\langle \mu s_n \exp(-H/kT) \rangle}{\langle \exp(-H/kT) \rangle} \tag{1.7}$$

where $\langle \ldots \rangle$ stands for a sum over configurations. M can be written

$$M = -\frac{1}{N} \frac{\partial}{\partial B} \left(\frac{F}{kT} \right). \tag{1.8}$$

The Ising model, in 2 or more dimensions, can display spontaneous magnetization. Let the lattice be in an external magnetic field B. If M remains different from zero when the external field is removed, there is spontaneous magnetization. For example, this occurs if at zero temperature the state of lowest energy of the system has all the spins aligned. We can expect $M = 1$ at zero temperature but $M = 0$ at infinite temperature, where thermal energy overwhelms the spin-spin interaction. Only for temperatures T less than a critical value T_c does spontaneous magnetization occur. Spontaneous magnetization is also closely related to the behavior of the spin-spin correlation function defined by,

$$\Gamma_n = \frac{\langle s_n s_0 \exp(-H/kT)\rangle}{\langle \exp(-H/kT)\rangle}. \tag{1.9}$$

At zero temperature when all the spins line up Γ_n is unity, independent of n. However, as $T \to \infty$ thermal excitations wash out the spin-spin correlations, so Γ_n is different from zero only at the origin. We will in fact see that the way Γ_n falls to zero as n increases will depend on whether $T < T_c$, $T = T_c$ or $T > T_c$. The Ising model, therefore, contains two lengths: the lattice spacing, and the distance over which spin-spin correlations are appreciable.

2.

In the first section we discussed the Ising model and the phenomenology of a critical point, such as the critical temperature at which spontaneous magnetization first occurs. We will study through seven sections the problem of critical behavior looking specifically at phenomena above the critical temperature where, although there is no spontaneous magnetization, the system feels the presence of the nearby critical point. In this section we will consider various theories of the critical point at the handwaving level: Landau's form of mean field theory, and the Kadanoff theory of effective interactions involving block spins. This second approach will provide the background for the renormalization group. (Some standard references on critical phenomena are refs. [95–101].)

2.1. *Elementary properties of systems near their critical temperature*

If we plot the spontaneous magnetization versus temperature for a ferromagnet, an antiferromagnet or the Ising model, one finds a curve similar to that shown in fig. 2.1. There is a critical temperature, T_c, at which spontaneous magnetization first occurs. The curve rises to some finite value at temperature zero. In the region just below the critical temperature, the magnetization is well approximated by a power law,

$$M \propto (T_c - T)^{\beta} \tag{2.1}$$

where β is an example of a critical exponent. Theories of critical behavior are mainly concerned with predicting such critical exponents or, at least, finding relations between them.

There is also the correlation length which we introduced earlier. Consider two spins, one at lattice site n, one at the origin and define the correlation function,

$$\Gamma_n = \langle s_n s_0 \exp(-H/kT)\rangle / \langle \exp(-H/kT)\rangle \tag{2.2}$$

where $\langle \ldots \rangle$ indicates a sum over all configurations. The correlation function has the property

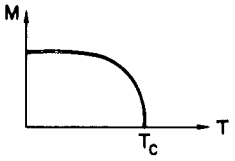

Fig. 2.1. A graph of magnetization versus temperature for a ferromagnet. T_c is the critical temperature.

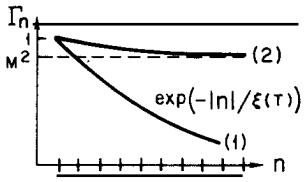

Fig. 2.2. Spin-spin correlation function versus distance between sites. Curve (1) is an example with $T > T_c$; curve (2) is an example with $T < T_c$.

that if $T > T_c$, then Γ_n falls off with $|n|$ (curve 1 in fig. 2.2). The asymptotic behavior for large $|n|$ is thought to be

$$\Gamma_n \sim \exp\{-|n|/\xi(T)\} \tag{2.3}$$

where $\xi(T)$ is the correlation length. (There are various ways of defining the correlation length. It is not evident that the qualitative definition of section 1 (see, in particular, the discussion p. 78) agrees with the precise definition of (2.3). They will in fact not always agree, but for the Ising model above T_c they do agree so far as is known.) Unfortunately, due to the complexity of the system, the validity of (2.3) has not been proved in general. For $T < T_c$, the correlation function is expected to approach the square of the magnetization divided by the square of the magnetic moment as n increases (curve 2, fig. 2.2). In other words, for large n actual correlation has disappeared and what remains is the product of the expectation values of the single spins. At the critical temperature the correlation function falls to zero for large n because there is no spontaneous magnetization. However, it falls very slowly (fig. 2.3) because we are at a transition point. The correlation function at T_c is expected to fall as a power of n and, by convention, is written in the form

$$\Gamma_n \sim 1/|n|^{d-2+\eta} \tag{2.4}$$

where d is the dimensionality of the system. In two dimensions we know that (2.4) is the correct form of the correlation function [95]. However, in three dimensions the statement is just a guess. η defines a second critical exponent.

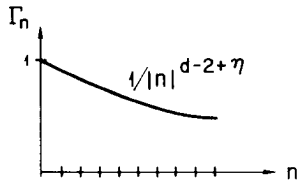

Fig. 2.3. Correlation function at the critical temperature.

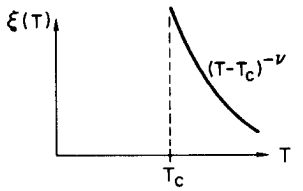

Fig. 2.4. Correlation length plotted against temperature.

One can also consider the correlation length $\xi(T)$ itself. $\xi(T)$ sets the scale for the falloff of the correlation function when T is above T_c. As shown in fig. 2.4, $\xi(T)$ approaches infinity as $T \to T_c$ from above:

$$\xi(T) \sim (T - T_c)^{-\nu}. \tag{2.5}$$

There is also a similar curve for the magnetic susceptibility,

$$\chi \equiv \left.\frac{\partial M}{\partial H}\right|_{H=0} \tag{2.6}$$

which behaves near the critical temperature as

$$\chi \sim (T - T_c)^{-\gamma}. \tag{2.7}$$

Similarly, the specific heat (at constant volume) C_v is controlled near the critical temperature by the critical exponent α,

$$C_v \sim (T - T_c)^{-\alpha}. \tag{2.8}$$

We collect in table 2.1 both experimental and theoretical values of the various exponents. The theoretical predictions come from the mean field (Landau) theory which will be described below,

Table 2.1

Critical exponent	Experiment	Mean field theory	Three-dimensional Ising model	Two-dimensional Ising model
β	0.3–0.38	$\frac{1}{2}$	0.31	0.125
η	0–0.1	0	0.056	0.25
ν	0.6–0.7	$\frac{1}{2}$	0.64	1.0
γ	1.2–1.4	1	1.25	1.75
α	0–0.1	0	0.12	0

the two-dimensional Ising model solved exactly by Onsager, and the three-dimensional Ising model. In the three-dimensional case the critical exponents are calculated approximately using the high temperature expansion of the model carried to very high order. To obtain critical exponents Padé approximant techniques are used. These assume power behavior near the critical point for the relevant thermodynamic quantities. The data come from ferromagnets, antiferromagnets, liquid-gas transitions, binary alloys and superfluid helium. Typically one or two critical exponents can be measured in each system. The table has been organized to illustrate the idea (to be explained in later sections) that the mean field theory accounts for systems with more than four dimensions, and that exponents depend continuously on dimensionality below 4 dimensions. One can plot critical exponents against the dimensionality of the system and construct a smooth interpolating curve between the various dimensions. Fig. 2.5 depicts such a plot for the exponent γ. The curve has a break in its derivative at dimension four and is smooth elsewhere.

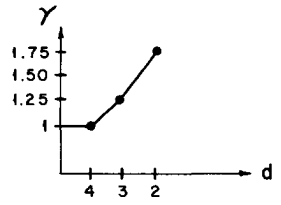

Fig. 2.5. Critical exponent γ plotted against the dimension of the physical system.

We will concentrate on the exponents η, ν and γ which have to do with the behaviour of the system at or above the critical temperature when no external field is present. We will not work in the two-phase region, although it would be more interesting; all the fundamental problems are present above T_c.

2.2. *The search for analyticity*

The various theories of critical phenomena are characterized by a "search for analyticity". The formulas we have written down, such as for spontaneous magnetization, are nonanalytic at the critical temperature. There is nothing really unusual about this, but it is difficult to see, starting from a fundamental formulation of the Ising model, for example, how such nonanalytic behavior arises. Recall that one calculates the thermodynamic functions from the partition function,

$$Z = \sum_{\{s\}} \exp\left\{K \sum_n \sum_{\hat{i}} s_n s_{n+\hat{i}}\right\}, \qquad K = J/kT. \tag{2.9}$$

As long as the number N of spins is finite, then Z is an analytic function of K. Such a system clearly cannot display critical behavior. However, once we take the thermodynamic limit of a spin system filling all of space, the analyticity is no longer guaranteed. As the volume V of the system tends to infinity, various sums will also become unbounded. The nonanalytic behavior we wish to discover here, however, is usually masked by these large volume effects. Thus, while nonanalyticity is allowed in principle, it has been very difficult to obtain the correct nonanalytic behavior in practice from expressions such as (2.9). Because of this one tries to find a description of critical behavior in terms of analytic functions, with the hope that the analytic functions could be obtained more easily by practical methods. The various theories can be characterized by their choice of analytic functions with which to describe critical phenomena. This is the "search for analyticity". We will discuss this problem at two levels: First, the mean field theory which makes a very simple theory of the analyticity; and the second, the renormalization group itself.

2.2.1. *Mean field theory*

A result of mean field theory is that there is a thermodynamic function which is analytic in its variables at T_c. Namely, the magnetic field H is an analytic function of the magnetization (per spin) M and the temperature T. This suggests Landau's hypothesis [102] that the thermodynamic properties of the system should be derivable from a free energy G which is an analytic function of M and T. Then H would be given by,

$$H = \frac{\partial}{\partial M} G(M, T). \tag{2.10}$$

What are the consequences of this approach? Near T_c where M is small we can write a Taylor series for G in powers of M,

$$G(M, T) = G_c(T) + r(T)M^2 + u(T)M^4 + \ldots \tag{2.11}$$

where M is the magnetization per unit spin. Only even powers of M appear in the expansion because of the up-down symmetry of the Ising lattice. The magnetic field is given by

$$H = 2r(T)M + 4u(T)M^3 + \ldots . \tag{2.12}$$

If we just consider the M and M^3 terms and assume that $u > 0$ (such that the magnetic field grows with M when M is large) we are left with two possible phenomena. If $r > 0$, then G will

have one minimum as a function of M (fig. 2.6a). But if $r < 0$, G can have two extra minima corresponding to nonvanishing magnetizations (fig. 2.6b) and hence three local extrema corresponding to $H = 0$. However, the system will choose a state of minimum free energy which corresponds, according to fig. 2.6b, to a state of nonvanishing magnetization. We see that the possibility $r > 0$ corresponds to $T > T_c$ where there is no spontaneous magnetization, while $r < 0$ means $T < T_c$.

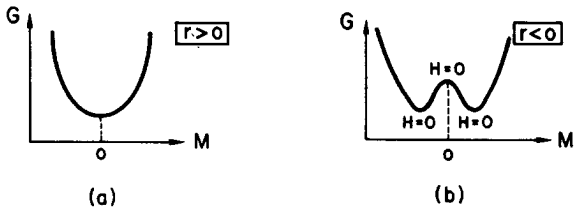

Fig. 2.6. Plots of free energy versus magnetization for the possible choices of the parameter r.

Continuing with these analyticity assumptions we can obtain some of the critical exponents. Namely, it is clear that $r(T_c) = 0$ while $u(T_c)$ will attain some particular value. The next stage is to assume that for T near T_c,

$$r(T) \propto (T - T_c). \tag{2.13}$$

Then at $H = 0$ it follows from (2.12) that,

$$M = \sqrt{-\frac{r(T)}{2u(T)}} \sim (T_c - T)^{1/2} \tag{2.14}$$

from which we identify

$$\beta = \tfrac{1}{2} \tag{2.15}$$

as the first critical exponent. If $T > T_c$, then near zero magnetization,

$$M = H/2r(T) \tag{2.16}$$

from which we identify the susceptibility

$$\chi = 1/2r(T) \sim (T - T_c)^{-1} \tag{2.17}$$

and another critical exponent, γ, is obtained:

$$\gamma = 1. \tag{2.18}$$

The values of the exponents ν and η can be obtained by extending these ideas to the case of space-dependent fields. One obtains $\nu = \frac{1}{2}$ and $\eta = 0$.

Unfortunately, as we have seen from table 2.1, the mean field theory does not predict the critical exponents very well. For example, accurate experiments with liquids, with $(T - T_c)/T_c$ as small as 10^{-5}, give $\beta = 0.35 \pm 0.01$ [96] as opposed to the mean field prediction of 0.5.

However, mean field theory provides a very simple picture of critical behavior. And, although its predictions are not very good, they are not very bad, either. In our later description of critical exponents the mean field theory results will act as zeroth order approximations. Then a systematic expansion in the dimension of the system will allow us to obtain more accurate results.

2.2.2. *Kadanoff theory*

Now we turn to the second level of discussion. We will discuss the Kadanoff picture of the critical point [103]. The basic results of the Kadanoff picture are relations among critical exponents which had been guessed earlier by equally heuristic arguments (see [95–101]). These relations will not be discussed here. What will be discussed is where one has analyticity, and how the critical singularities are generated starting from analytic functions.

We have used considerable poetic license in describing Kadanoff's picture. The following paragraphs explain the spirit of his ideas but details of his work will be considerably distorted.

Kadanoff is concerned explicitly with the problem of the large correlation length near T_c. A direct calculation of critical behavior requires that one consider (at the minimum) all the spins in a volume the size of the correlation length – a hopeless task. In contrast, away from T_c where the correlation length is small, only a few spins need be considered at any one time. Standard techniques are available for such problems, such as Feynman graphs and perturbation theory.

Kadanoff had a brilliant idea which allows the hopeless problem with a large correlation length to be replaced by one with a small correlation length. The idea will be explained for a plane lattice, for simplicity. Consider a small region containing four spins, say. Because the correlation length is so long near the critical temperature, all the spins in such a little block should be strongly correlated. Kadanoff supposed that they are so well correlated that the four spins in one block have only two possible states: all spins up or all spins down. This means that the block of four spins acts like a single effective spin. Now suppose ξ is 1000 in units of the lattice spacing (fig. 2.7a). Group the spins into blocks of four spins each (fig. 2.7b). Each block is supposed to have only two spin degrees of freedom, so there is a single spin variable for each block. Therefore, one can replace the original lattice with an effective lattice where now $\xi = 500$ in units of the effective lattice spacing (fig. 2.7c). In this way the problem with $\xi = 1000$ is reduced to a problem with $\xi = 500$. Repetition of this analysis allows further reductions in ξ, to

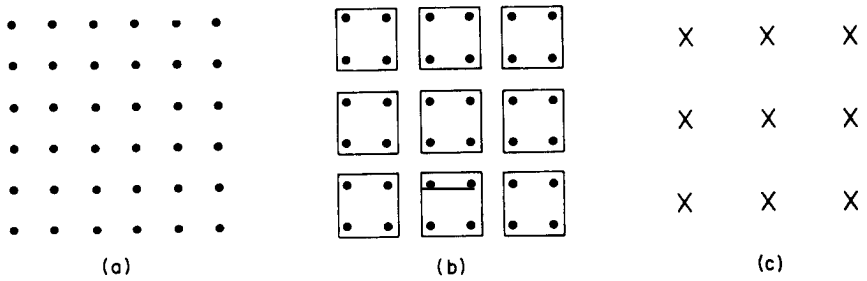

Fig. 2.7. Visualization of Kadanoff's construction of block spins. The original lattice (a) of spins at each site is divided into blocks (b) of 4 spins per block which is replaced by a new lattice (c) of "effective" spins.

250, 125, etc., until finally one has an effective theory with $\xi \sim 1$. This will be discussed further below. If the original spins have only nearest neighbor couplings, then the effective block spins also can have only nearest neighbor interactions (see later). It is convenient to define renormalized block spins such that their magnitude is ± 1 instead of ± 4. Then the energy/kT of the block spins is

$$\sum_{n,i} K_1 s_n^{(1)} s_{n+i}^{(1)} \tag{2.19}$$

where K_1 is a constant and $s_n^{(1)}$ is the block spin variable and n labels sites on the effective lattice

(X's in fig. 2.7c). The only practical difference between Kadanoff's effective interaction (2.19) and the original interaction $\Sigma_{n,i} K s_n s_{n+\hat{i}}$ is the change in constants from K to K_1. It is easy to determine the constant K_1. There are two spin-spin interactions from the original interaction which couple the block spin $s_n^{(1)}$ to the block spin $s_{n+\hat{i}}^{(1)}$. If these block spins are parallel, the coupling energy is $2K$; if they are antiparallel the coupling energy is $-2K$. So

$$K_1 = 2K. \tag{2.20}$$

Suppose now one solves the original interaction and obtains the correlation length $\xi(K)$ as a function of K, in units of the original lattice spacing. The correlation length for the lattice of block spins will be $\xi(K_1)$, in units of the block spin lattice spacing, since the block-spin interaction has Ising form. But now the correlation length in units of block-spin spacing must be $\frac{1}{2}$ of the original correlation length, i.e.

$$\xi(K_1) = \tfrac{1}{2}\,\xi(K). \tag{2.21}$$

Given that $\xi(K_1) = \frac{1}{2}\,\xi(K)$ whenever $K_1 = 2K$, the dependence of $\xi(K)$ on K is severely restricted. Suppose one is at the critical temperature. There is a corresponding value $K_c = J/kT_c$ for K. At K_c, ξ is infinite. But now the correlation length for the block spins must also be infinite; this means $\xi(K_1) = \infty$ which is possible only if K_1 is also K_c. So

$$K_c = K_1 = 2K_c, \tag{2.22}$$

which gives $K_c = 0$ or $K_c = \infty$, not $K_c = J/kT_c$.

This is a nonsensical result: K_c is a finite number, not 0 or ∞. The trouble originates with the assumption that the spins in a block align exactly. Kadanoff did not actually assume this. He proposed only that the block would behave as if it had only two possible states, and therefore could be replaced by an effective block spin. However these two states would not be the states with all spins up or down. Kadanoff proposed that there would be an effective Ising interaction for the block spins, with K_1 being some function $f(K)$, but the function f would be more complicated than $2K$ obtained above: Kadanoff gave no prescription for determining $f(K)$. What Kadanoff does assume is that $f(K)$ is still an analytic function of K, even for $K = K_c$. The rationale for this is the hope that only the spins in the immediate neighborhood of the block n affect the calculation of K_1, even though this calculation cannot be spelled out. Nonanalyticity at K_c should occur only for quantities involving the entire lattice.

The statement of the Kadanoff assumptions is that there exists an analytic function $f(K)$ such that

$$\xi[f(K)] = \tfrac{1}{2}\,\xi(K). \tag{2.23}$$

What does this imply for critical behavior? First one must have

$$K_c = f(K_c) \tag{2.24}$$

so that $\xi[f(K_c)]$ is infinite. Secondly, suppose K is near K_c. Approximately, one can write

$$f(K) = f(K_c) + \lambda(K - K_c) \tag{2.25}$$

where $\lambda = \mathrm{d}f/\mathrm{d}K$ for $K = K_c$. This means

$$f(K) - K_c = \lambda(K - K_c). \tag{2.26}$$

Now suppose $\xi(K)$ behaves as $(K - K_c)^{-\nu}$ for K near K_c. Then one must have

$$\frac{\xi[f(K)]}{\xi(K)} = \left(\frac{f(K) - K_c}{K - K_c}\right)^{-\nu} \tag{2.27}$$

if K is near K_c. But from (2.23) and (2.26), this equation reduces to

$$\tfrac{1}{2} = [\lambda]^{-\nu}. \tag{2.28}$$

If one could determine $f'(K_c)$, one could calculate ν from this equation:

$$\nu = \ln 2/\ln \lambda. \tag{2.29}$$

For example, if $f(K)$ were the explicit form $2K$ obtained earlier, then λ would be 2 and ν would be 1.

Equation (2.26) does not force $\xi(K)$ to have power law behavior for $K \to K_c$; the most general form for $\xi(K)$ consistent with (2.23) and (2.26) is

$$\xi(K) = (K - K_c)^{-\nu} F[\ln(K - K_c)], \tag{2.30}$$

where $F(x)$ is a periodic function of x with period $\ln \lambda$. The reason for the periodicity is that if $\xi(K)$ is known, then $\xi(K_1)$ is also known where $K_1 - K_c = \lambda(K - K_c)$; changing $K - K_c$ by a factor λ is equivalent to translating $\ln(K - K_c)$ by an amount $\ln \lambda$. Because F is periodic, the behavior of $\xi(K)$ is not qualitatively different from $(K - K_c)^{-\nu}$.

To determine $\xi(K)$ precisely one can imagine proceeding as follows. Suppose $\xi(K)$ is known for $K > 2K_c$. For K above $2K_c$, ξ should not be very large and should be relatively easy to calculate. Now suppose K lies in the range $K_c < K < 2K_c$. Construct a sequence of effective constants K_1, K_2, etc. satisfying

$$K_{n+1} = f(K_n) \tag{2.31}$$

with $K_1 = f(K)$. Then Kadanoff's formula implies

$$\xi(K) = 2^n \xi(K_n). \tag{2.32}$$

This means in particular that no matter what value K takes ($> K_c$ itself), one can choose n so large that $\xi(K_n)$ is small enough that $K_n > 2K_c$. Then $\xi(K_n)$ is known and $\xi(K)$ is determined by (2.32). This calculation is of course possible only if the function $f(K)$ is known.

What is important in Kadanoff's analysis is the idea that starting from an analytic function $f(K)$ one generates a nonanalytic behavior for $\xi(K)$ at the point K_c for which $f(K_c) = K_c$. Furthermore one does *not* get an explicit value for ν independent of the function $f(K)$; to determine ν one must know the function $f(K)$. Hence ν need not be the mean field value $\frac{1}{2}$. In fact ν can be irrational, contrary to the dreams of some statistical mechanicians.

In the following, Kadanoff's idea that there exist effective block-spin interactions with coupling constants analytic in T will be realized in various forms, with all functions given explicitly. Critical exponents such as ν will be computed explicitly.

3. Trivial example of the renormalization group: The Gaussian model

In this section we will begin to make Kadanoff's intuitive ideas quantitative. The relationship between the block spin interaction and the original interaction will be worked out and a critical exponent (ν) will be computed for a trivial model – the Gaussian model [104] – to illustrate the ideas involved. In this case the exponent ν has the mean field value $\frac{1}{2}$.

The Gaussian model can be obtained by modifying the Ising model. First one writes the partition function in terms of integrals. Namely,

$$Z = \prod_m \int_{-\infty}^{\infty} ds_m 2\delta(s_m^2 - 1) \exp\left\{K \sum_n \sum_i s_n s_{n+\hat{i}}\right\}. \tag{3.1}$$

This is a trivial rewriting of the original partition function. Now imagine smoothing the delta function (fig. 3.1a) to a smooth distribution around $s_n = \pm 1$ (fig. 3.1b), to a smooth Gaussian

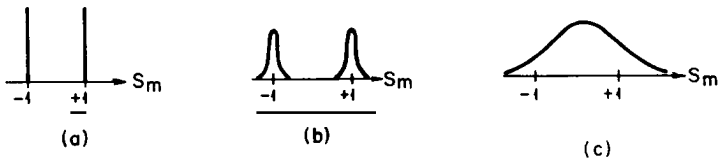

Fig. 3.1. The transition from the Ising to the Gaussian model. The Ising model (a) has spin up or spin down at each lattice site. Model (b) has spin variables which peak about the Ising values. The Gaussian model (c) has spin variables at each site with smooth, Gaussian distributions about zero

(fig. 3.1c). Of course fig. 3.1c bears little resemblance to the original Ising model. If we make this replacement in the partition function anyway,

$$Z = \prod_m \int_{-\infty}^{\infty} ds_m \exp\left\{-\tfrac{1}{2} b s_m^2\right\} \exp\left\{K \sum_n \sum_i s_n s_{n+\hat{i}}\right\} \tag{3.2}$$

where b is an arbitrary constant. This formula defines the Gaussian model [104]. Later we will allow the generalization,

$$\exp\left\{-\tfrac{1}{2} b s_n^2\right\} \to \exp\left\{-\tfrac{1}{2} b s_n^2 - u s_n^4\right\} \tag{3.3}$$

where u is a positive number. The case u near 1 (not small) can be discussed using the recursion formula to be derived later. In this case the smooth model begins to approach the real Ising model. Finally, if $u \to \infty$ and $b \to -\infty$, with $b = -4u$, the original Ising model is recovered (if one includes a constant factor $(u/\pi)^{1/2} \exp(-u)$ per spin).

Let us review some characteristics of Gaussian integrals. It is convenient to introduce matrix notation,

$$\tilde{s} A s \equiv \sum_n \sum_m s_n A_{nm} s_m, \qquad \tilde{\rho} s \equiv \sum_n \rho_n s_n \tag{3.4}$$

with A assumed to be symmetric for $n \leftrightarrow m$. Then, using techniques to be developed shortly, the following integral can be evaluated,

$$\prod_n \int_{-\infty}^{\infty} ds_n \exp\left\{-\tfrac{1}{2} \tilde{s} A s + \tilde{\rho} s\right\} = C \exp\left\{\tfrac{1}{2} \tilde{\rho} A^{-1} \rho\right\} \tag{3.5}$$

where C, a normalization factor, is a function of A. (Specifically $C = (\det A)^{-1/2} (2\pi)^{N/2}$ where N is the number of sites.) Eq. (3.5) is derived by translating s_n in order to complete the square in the

integrand. (See eq. (3.21) below for an example of a translation.) From eq. (3.5) one can derive all integrals of the form

$$\prod_n \int_{-\infty}^{\infty} ds_n (s_{m_1} s_{m_2} \ldots s_{m_k}) \exp\left\{-(\tilde{s} A s)/2\right\} = I(m_1 \ldots m_k). \tag{3.6}$$

These integrals will be needed in section 4. The procedure is to differentiate eq. (3.5) with respect to $\rho_{m_1} \ldots \rho_{m_k}$. This gives

$$\prod_n \int_{-\infty}^{\infty} ds_n (s_{m_1} \ldots s_{m_k}) \exp\left\{-\tfrac{1}{2} \tilde{s} A s + \tilde{\rho} s\right\} = \frac{\partial}{\partial \rho_{m_1}} \ldots \frac{\partial}{\partial \rho_{m_k}} C \exp\left\{\tfrac{1}{2} \tilde{\rho} A^{-1} \rho\right\}.$$

The second step is to set $\rho_m = 0$ for all m. If k is odd the integral vanishes; if k is even one needs to consider only the term of order ρ^k from the exponential, giving

$$I(m_1, \ldots, m_k) = \frac{\partial}{\partial \rho_{m_1}} \ldots \frac{\partial}{\partial \rho_{m_k}} C \left\{\tfrac{1}{2} \tilde{\rho} A^{-1} \rho\right\}^{k/2} / (k/2)!. \tag{3.7}$$

One now has k derivatives applied to $k/2$ factors of $(\frac{1}{2} \tilde{\rho} A \rho)$. To illustrate the calculation of the derivatives consider the case $k = 4$. Then one has to calculate

$$\frac{\partial}{\partial \rho_{m_1}} \frac{\partial}{\partial \rho_{m_2}} \frac{\partial}{\partial \rho_{m_3}} \frac{\partial}{\partial \rho_{m_4}} \left\{\tfrac{1}{2} \tilde{\rho} A^{-1} \rho\right\} \left\{\tfrac{1}{2} \tilde{\rho} A^{-1} \rho\right\} / 2!.$$

Differentiating yields a number of terms such as $(A^{-1})_{m_1 m_2} (A^{-1})_{m_3 m_4}/8$. The total number of such terms is 24, namely the total number of ways of pairing each derivative with a ρ or $\tilde{\rho}$. However, many of these terms are identical to the one just cited. Due to the symmetry of A^{-1}, the term $(A^{-1})_{m_2 m_1} (A^{-1})_{m_3 m_4}/8$ is identical. Obviously the term $(A^{-1})_{m_3 m_4} (A^{-1})_{m_1 m_2}/8$ is identical. Combining all identical terms one finds

$$I(m_1, \ldots, m_4) = C\{(A^{-1})_{m_1 m_2} (A^{-1})_{m_3 m_4} + (A^{-1})_{m_1 m_3} (A^{-1})_{m_2 m_4} + (A^{-1})_{m_1 m_4} (A^{-1})_{m_2 m_3}\}. \tag{3.8}$$

The general rule for any k is constructed similarly. The result is as follows. Let $\overbrace{s_{m_1} s_{m_2}}$ stand for $(A^{-1})_{m_1 m_2}$. One calls $\overbrace{s_{m_1} s_{m_2}}$ a "contraction". Then $I(m_1, \ldots, m_k)$ is the sum of all possible ways of contracting $s_{m_1} s_{m_2} \ldots s_{m_k}$ in pairs, such that all s_{m_i} are in a contraction. For example

$$I(m_1, m_2, m_3, m_4) = C\left\{\overbrace{s_{m_1} s_{m_2}}\ \overbrace{s_{m_3} s_{m_4}} + \overbrace{s_{m_1} s_{m_3}}\ \overbrace{s_{m_2} s_{m_4}} + \overbrace{s_{m_1} s_{m_4}}\ \overbrace{s_{m_2} s_{m_3}}\right\} \tag{3.9}$$

which is the same as eq. (3.8). $I(m_1 \ldots m_6)$ consists of $\overbrace{s_{m_1} s_{m_2}}\ \overbrace{s_{m_3} s_{m_4}}\ \overbrace{s_{m_5} s_{m_6}}$ plus 14 other terms. Note that it is not necessary for the m_i to be distinct. For example, suppose one is dealing with trivial 1 x 1 matrices so that the m_i are necessarily all equal to 1. Then the integral (3.6) of s_1^4 is $3CA_{11}^{-2}$, the integral of s_1^6 is $15CA_{11}^{-3}$, etc., in agreement with standard formulae. The rule for $I(m_1, \ldots, m_k)$ is analogous to Wick's theorem in quantum field theory [105].

In the cases of interest $A_{n,m}$ only depends on the difference $n - m$ and the inverse matrix $(A^{-1})_{n,m}$ can be computed by Fourier transforms: write

$$A_{n-m} = \int_q \exp\left\{iq \cdot (n - m)\right\} A(q) \tag{3.10}$$

where

$$\int_q \equiv \frac{1}{(2\pi)^d} \int_{-\pi}^{\pi} dq_1 \ldots \int_{-\pi}^{\pi} dq_d.$$

The inverse matrix is obtained by taking the transform of the ordinary inverse of $A(q)$,

$$(A^{-1})_{n-m} = \int_q \exp\{iq \cdot (n-m)\} A^{-1}(q) \tag{3.11}$$

since A is diagonal in the q representation.

Return now to the Gaussian model. We can rewrite the expression of interest,

$$K\sum_n \sum_i s_n s_{n+\hat{i}} - \tfrac{1}{2} b \sum_n s_n^2 = -\tfrac{1}{2} K \sum_n \sum_i (s_{n+\hat{i}} - s_n)^2 - (\tfrac{1}{2} b - dK) \sum_n s_n^2 \tag{3.12}$$

where we are supposing for convenience that the sum over i runs only over positive values. d is the dimensionality of the lattice. When written in terms of σ_q, $\sigma_q = \Sigma_n \exp\{-iq \cdot n\} s_n$, (3.12) becomes

$$-\tfrac{1}{2} \int_q \left\{K \sum_i |\exp(iq_i) - 1|^2 + \tilde{r}\right\} \sigma_q \sigma_{-q} \tag{3.13}$$

with

$$\tilde{r} = b - 2dK \tag{3.14}$$

where q_i is the ith component of q. There are now three further changes that will be made on the model. The first change is to replace $|\exp(iq_i) - 1|^2$ by its form for small q, namely q_i^2. The second is to change the range of integration over q to be $0 < |q| < 1$ instead of $-\pi < q_i < \pi$. The third change is to renormalize the spins so that $K = 1$. So finally, the interaction is

$$-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q \sigma_{-q} \tag{3.15}$$

with $\int_q$ being $(2\pi)^{-d} \int d^d q$ for $0 < |q| < 1$ and $r = \tilde{r}/K$.

The second change causes a conceptual difficulty. With $0 < |q| < 1$, one cannot relate the functional variable σ_q to the ordinary variables s_n. Therefore, one is forced to define the partition function as a functional integral over σ_q rather than ordinary integrals over the s_n. However, for present purposes this is not a difficulty. The reason is that the method of translating the integration variable is valid for functional integrals as well as ordinary integrals, and is as easy to use. The reason one can translate variables in a functional integral is that a functional integral is generally defined as a limit of ordinary integrals, each of which can be translated. For example, the functional integral over σ_q can be described as a limit of ordinary integrals over variables σ_{q_m} in the limit that the spacing of the points q_m goes to zero (see, e.g., [106]).

The replacement of $|\exp(iq_i) - 1|^2$ by q_i^2 is not an essential change in the model because we will ultimately be interested only in the long wavelength behavior of the model which comes from the part with small q. Likewise the change in the range of $|q|$ is inessential.

The correlation function of the model will now be calculated. It is appropriate to introduce a spin field $s(x)$:

$$s(x) = \int_q \exp(iq \cdot x)\, \sigma_q \qquad (0 < |q| < 1) \tag{3.16}$$

to replace s_n. Then the correlation function is

$$\Gamma(x) = \int_\sigma s(x) s(0) \exp\left\{-\tfrac{1}{2} \int_q \sigma_q \sigma_{-q} (q^2 + r)\right\} \Big/ Z \tag{3.17}$$

where $\int_\sigma$ denotes the functional integral over functions $\sigma(q)$ and Z is the functional integral of the exponential alone. The result for $\Gamma(x)$ is

$$\Gamma(x) = \int_q \exp(iq \cdot x) \frac{1}{q^2 + r} = \int_q \exp(iq \cdot x)\, \Gamma_q . \tag{3.18}$$

This result will now be derived. One has

$$\Gamma(x) = \int_q \int_{q_1} \exp(iq \cdot x) \int_\sigma \sigma_q \sigma_{q_1} \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q \sigma_{-q}\right\} \Big/ Z. \tag{3.19}$$

The functional integral is computed from the generating function

$$Z(j) = \frac{1}{Z} \int_\sigma \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q \sigma_{-q} + \int_q j_{-q} \sigma_q\right\}. \tag{3.20}$$

Translating variables means in this case writing

$$\sigma_q = \frac{j_q}{q^2 + r} + \sigma'_q . \tag{3.21}$$

In terms of σ'_q,

$$Z(j) = \frac{1}{Z} \int_{\sigma'} \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma'_q \sigma'_{-q} + \tfrac{1}{2} \int_q j_q j_{-q} \left(\frac{1}{q^2 + r}\right)\right\}. \tag{3.22}$$

The functional integral is proportional to Z, giving

$$Z(j) = \exp\left\{\tfrac{1}{2} \int_q j_q j_{-q} \left(\frac{1}{q^2 + r}\right)\right\}. \tag{3.23}$$

Expanding (3.20) and (3.23) to second order in j one sees that

$$\tfrac{1}{2} \int_q \int_{q_1} j_{-q} j_{-q_1} \int_\sigma \sigma_q \sigma_{q_1} \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q \sigma_{-q}\right\} \Big/ Z = \tfrac{1}{2} \int_q j_q j_{-q} \frac{1}{q^2 + r}. \tag{3.24}$$

This equation must be an identity in the function j, which means

$$\int_\sigma \sigma_q \sigma_{q_1} \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q \sigma_{-q}\right\} \Big/ Z = \delta(q + q_1)/(q^2 + r) \tag{3.25}$$

where $\delta(q + q_1)$ is shorthand for $(2\pi)^d \delta^d(q + q_1)$. Substitution of this formula in (3.19) gives the result (3.18) for $\Gamma(x)$.

The correlation length ξ is customarily defined in terms of the behavior of $\Gamma(x)$ for $|x| \to \infty$. If the range of q were infinite, the behavior of $\Gamma(x)$ for large x would be governed by the singularity in $(q^2 + r)^{-1}$ for $q = \pm i\sqrt{r}$, resulting in

$$\Gamma(x) \propto \exp(-\sqrt{r}|x|) \tag{3.26}$$

for large $|x|$ (apart from a power of $|x|$ in front of the exponential). From this one gets $\xi = 1/\sqrt{r}$. Unfortunately, the sharp boundary at $|q| = 1$ leads to another term in $\Gamma(x)$ for large x, behaving as $\cos(|x|)$. This latter term is an artifact of the model; it can be removed by averaging over a large but finite region in x space. When ξ is large (r small) the term $\exp(-\sqrt{r}|x|)$ will not be

affected by this averaging. However a more convenient procedure for our purposes is to introduce an alternative definition of the correlation length. The correlation length as originally defined is determined by the leading singularity of Γ_q: if the leading singularity of Γ_q is at $q = \pm iq_s$ then ξ is $1/q_s$. The location of this singularity can be determined roughly by comparing the derivative $d\Gamma_q/dq^2$ with Γ_q itself when $q = 0$, namely,

$$\xi^2 \propto -\left.\frac{d\Gamma_q/dq^2}{\Gamma_q}\right|_{q=0} \tag{3.27}$$

Equivalently,

$$\xi^2 = \int x^2 \Gamma(x)\, d^d x \Big/ \int \Gamma(x)\, d^d x.$$

With this definition, ξ is called the "effective range of correlation" [107]. For $\Gamma_q = 1/(q^2 + r)$ this formula gives

$$\xi^2 \propto 1/r \tag{3.28}$$

and hence $\xi \propto 1/\sqrt{r}$ as before.

In the Gaussian model r is linear in the temperature T. To be precise, with $K = J/kT$ (from eq. (2.9)) one has

$$r = (bkT - 2dJ)/J \tag{3.29}$$

(see eqs. (3.14) and (3.15)).

The critical point corresponds to $r = 0$ and $\xi = \infty$; clearly $\xi \propto (T - T_c)^{-1/2}$ for $T > T_c$. Hence, the Gaussian model gives the mean field value $\frac{1}{2}$ for ν. Also this gives $kT_c = 2dJ/b$. The critical value of K is $K_c = b/2d$.

An exact formulation of the renormalization group can be defined for the Gaussian model; it illustrates the ideas that will be applied later to non-trivial models.

As a first step in the renormalization group transformation, let us integrate out the spin components σ_q with $\frac{1}{2} < |q| < 1$, leaving unintegrated the components σ_q with $0 < |q| < \frac{1}{2}$. In other words one integrates out the rapidly fluctuation parts of the spin field $s(x)$, leaving unintegrated the slowly fluctuating parts. This is a way of realizing in spirit Kadanoff's idea of producing an effective interaction involving only block spin variables; namely, one integrates out the other variables orthogonal to the block spin variables. In our case one thinks of the long wavelength spin components σ_q for $|q| < \frac{1}{2}$ as analogous to Kadanoff's block spin variables.

The integration of σ_q for $\frac{1}{2} < |q| < 1$ will be done in the functional integral for the partition function, as opposed to the functional integral for the correlation function. However the same calculation applies also to the correlation function integral if one considers the functional integral only for Γ_{q_1} for $|q_1| < \frac{1}{2}$, where

$$\delta(q_1 + q_2)\,\Gamma_{q_1} = \int_\sigma \sigma_{q_1}\, \sigma_{q_2}\, \exp\left\{-\tfrac{1}{2} \int_q (q^2 + r)\, \sigma_q\, \sigma_{-q}\right\} \Big/ Z. \tag{3.30}$$

If $|q_1|$ and $|q_2|$ are both less than $\frac{1}{2}$, then the explicit σ_{q_1} and σ_{q_2} terms are not involved in the calculation of integrals for $\frac{1}{2} < |q| < 1$.

The calculation of the first step defined above is trivial. The reason is that there is no coupling between σ_q for $|q| > \frac{1}{2}$ and σ_q for $|q| < \frac{1}{2}$. Hence the only result of the functional integral over σ_q with $|q| > \frac{1}{2}$ is a constant which multiplies $\exp\{-\frac{1}{2} \int_q' (q^2 + r)\, \sigma_q\, \sigma_{-q}\}$ where $\int_q'$ means an integral

over $0 < |q| < \frac{1}{2}$ only. Since this multiplicative constant cancels in the ratio of functional integrals defining Γ_q, the constant can be ignored. Thus the effective interaction produced by eliminating the high momentum spin components is simply $-\frac{1}{2}\int_q' (q^2 + r)\,\sigma_q\,\sigma_{-q}$.

The second step in constructing the renormalization group transformation is to perform some scale changes designed to make the effective interaction look as much like the original interaction as possible. Two scale changes will be introduced. A scale change in the momentum, namely $q' = 2q$, is made so that the new momentum variable q' has the range $0 < |q'| < 1$ like the original momentum q in the original interaction. The second scale change is to scale the spin variable itself:

$$\sigma_q = \zeta\,\sigma'_{q'} = \zeta\,\sigma'_{2q}. \tag{3.31}$$

The scale factor ζ will be determined later. In terms of the new spin variable σ', the effective interaction is

$$\mathcal{H}' = -\tfrac{1}{2}(\zeta^2 2^{-d}) \int_{q'} (q'^2/4 + r)\,\sigma'_{q'}\,\sigma'_{-q'}. \tag{3.32}$$

Now choose ζ so that the q'^2 term has the same coefficient as the q^2 term of the original interaction; this means

$$\zeta = 2^{1+d/2} \tag{3.33}$$

and

$$\mathcal{H}' = -\tfrac{1}{2}\int_{q'} (q'^2 + r')\,\sigma'_{q'}\,\sigma'_{-q'} \tag{3.34}$$

with

$$r' = 4r. \tag{3.35}$$

The transformation that takes

$$\mathcal{H} = -\tfrac{1}{2}\int_q (q^2 + r)\,\sigma_q\,\sigma_{-q} \tag{3.36}$$

into $\mathcal{H}'$ is an example of a renormalization group transformation.

The two scale changes introduced above are analogous to scale changes that occur in the Kadanoff picture. When one discusses the block spin interaction, one uses the spacing of blocks as a length scale instead of the original lattice spacing. This is analogous to the momentum scale change introduced here. Secondly, the magnitude of the block spins was rescaled so it would be ± 1 like the original spins; analogously, the spins σ_q are rescaled, although now the purpose is to make a particular term in the interaction have the same form as in the original interaction.

What is $\mathcal{H}'$ good for? It can be used to compute the correlation function Γ_q, provided $|q| < \frac{1}{2}$. The correlation function naturally defined for $\mathcal{H}'$ is a correlation function involving σ' instead of σ, namely,

$$\delta(q'_1 + q'_2)\,\Gamma'_{q'_1} = \int_{\sigma'} \sigma'_{q'_1}\,\sigma'_{q'_2}\,\exp\{\mathcal{H}'\}/Z'. \tag{3.37}$$

The scaling relation between σ and σ' gives

$$\delta(q_1 + q_2)\,\Gamma_{q_1} = \zeta^2\,\Gamma'_{2q_1}\,\delta(2q_1 + 2q_2). \tag{3.38}$$

Now $\delta(2q) = 2^{-d}\delta(q)$ so

$$\Gamma_q = \zeta^2 2^{-d}\Gamma'_{2q} = 4\Gamma'_{2q}. \tag{3.39}$$

Thus, if one computes Γ'_q from $\mathcal{H}'$, one can reconstruct $\Gamma_{q/2}$ from the scaling formula.

Now look specifically at the correlation length. The correlation length for $\mathcal{H}'$ is (using the modified definition),

$$\xi' = 1/\sqrt{r'}. \tag{3.40}$$

Because of the change in momentum scale in defining $\mathcal{H}'$, ξ' will be different from ξ. Since momenta were doubled in the transformation from $\mathcal{H}$ to $\mathcal{H}'$, lengths will be halved, and therefore ξ' should be $\xi/2$. This is indeed the case, since

$$\xi' = (r')^{-1/2} = (4r)^{-1/2} = \xi/2. \tag{3.41}$$

Here we have an explicit example of the relation between ξ' calculated from the effective interaction $\mathcal{H}'$ and ξ calculated for the original interaction. There are three relations involved. First there is the relation between ξ' and ξ: $\xi' = \frac{1}{2}\xi$. This relation is a consequence of the explicit change of scale in momentum introduced in the renormalization group transformation. This relation will be true of all models so long as we stick to the program of integrating out the momentum components σ_q with $|q| > \frac{1}{2}$ and then rescale the remaining momenta. The second relation is $r' = 4r$, the relation between the constants appearing in $\mathcal{H}$ and $\mathcal{H}'$. This relation is special to the Gaussian model. However, note that r' is an analytic function of r, even at the critical point. This analyticity will be true of all subsequent examples. Finally there is the dependence of ξ on r: $\xi = 1/\sqrt{r}$. Since $\mathcal{H}'$ has the same form as $\mathcal{H}$, the new correlation length ξ' has the same dependence on r': $\xi' = 1/\sqrt{r'}$. These relations are special to the Gaussian model; but it will always be true that if $\xi = X(r)$ then $\xi' = X(r')$ with the same function X.

The critical singularity of ξ can be calculated by the method of section 2. The variable r substitutes for the variable K of section 2. The function $f(K)$ of section 2 is now known explicitly: $r' = f(r) = 4r$. The critical value of r is $r_c = 0$, which satisfies $f(r_c) = r_c$. The parameter λ defined in section 2 is $df/dr = 4$. Thus, from eq. (2.29)

$$\nu = \ln 2/\ln \lambda = \tfrac{1}{2}. \tag{3.42}$$

In this case $\xi(r)$ must behave as $\xi(r) \propto (r - r_c)^{-1/2}$ without the extra periodic function $F[\ln(r - r_c)]$ of section 2. The reason is that the period of F is determined by the fact that ξ changes by a factor $2[\xi(r') = \frac{1}{2}\xi(r)]$. One could have chosen to make ξ change by an arbitrary factor s, by integrating σ_q for $s^{-1} < |q| < 1$ instead of $\frac{1}{2} < |q| < 1$. The result would still be $\nu = \frac{1}{2}$ but the extra function would have period $\ln s$ instead of $\ln 2$. This gives a contradiction unless F is a constant.

Thus, if $f(r)$ is known, one can determine both the critical point $r_c = 0$ ($K_c = b/2d$) and the critical singularity of ξ.

One starts with an analytic formula for $r' = 4r$ and obtains a nonanalytic form for ξ. Why? The reason is that as $r \to 0$, r' comes close to r but ξ' cannot come close to ξ, since ξ' must be $\xi/2$. It is this conflict that forces ξ to be singular when $r' = r = 0$.

4. The s^4 model

In this section we will generalize the discussion of the Gaussian model. In particular, we will add to the interaction a small quartic term. The renormalization problem now becomes nontrivial

but much more instructive. A perturbative definition of the renormalization group will be given and a nontrivial fixed point solution to the equations will be found. The discussion will also introduce the use of continuous, nonintegral spatial dimensions.

The new model is specified by the partition function [108–111]

$$Z = \int_{\sigma} \exp\{\mathcal{H}[\sigma]\} \tag{4.1}$$

where

$$\mathcal{H}[\sigma] = -\tfrac{1}{2} \int_{q} (q^2 + r)\, \sigma_q \sigma_{-q} - u \int_{q_1} \int_{q_2} \int_{q_3} \sigma_{q_1} \sigma_{q_2} \sigma_{q_3} \sigma_{-q_1-q_2-q_3}\,. \tag{4.2}$$

The first term in the Hamiltonian (to be called $\mathcal{H}_F$) is familiar from before. The parameter u is presumed small compared to one. This will allow us to describe the new term perturbatively. Also note that the new term is just

$$\mathcal{H}_I = -u \int_{x} s^4(x)$$

in configuration space.

Following the ideas of the previous section we attempt to define a new physical system in which the high frequency modes of the present system are integrated out. The effective Hamiltonian governing the new physical system will be defined to be as similar as possible to (4.2). This will involve considerable simplification which will be justified in the next section. In place of the constants r and u, new constants r' and u' will appear in the effective Hamiltonian. They will be determined by r and u via simple formulae. The procedure is similar to Anderson's treatment of the Kondo problem [112].

The construction proceeds as follows. Let the momentum integrals in the original system run over the range $0 < |q| < 1$. The new system is obtained, as in section 3, by integrating out the high frequency modes, $\frac{1}{2} < |q| < 1$. Write the function σ_q in the form,

$$\sigma_q = \sigma_{0,q} + \sigma_{1,q} \tag{4.3}$$

where

$$\begin{aligned} &\sigma_{1,q} = \sigma_q \quad \text{if} \quad \tfrac{1}{2} < |q| < 1, \text{ zero otherwise,} \\ &\sigma_{0,q} = \sigma_q \quad \text{if} \quad 0 < |q| < \tfrac{1}{2}, \text{ zero otherwise.} \end{aligned} \tag{4.4}$$

(This is the reverse of the notation of [111].) The functional integral Z then can be written as,

$$Z = \int_{\sigma} \exp\{\mathcal{H}[\sigma]\} = \int_{\sigma_0} \left\{ \int_{\sigma_1} \exp\{\mathcal{H}[\sigma_0 + \sigma_1]\} \right\}. \tag{4.5}$$

Integrate out the high frequency modes σ_1. This leaves,

$$Z = \int_{\sigma'} \exp\{\mathcal{H}'[\sigma']\} \tag{4.6}$$

where the new Hamiltonian $\mathcal{H}'$ is defined by,

$$\begin{aligned} &\exp\{\mathcal{H}'[\sigma']\} = \int_{\sigma_1} \exp\{\mathcal{H}[\sigma_0 + \sigma_1]\} \\ &\sigma_{0,q} = \zeta \sigma'_{(2q)}, \qquad 0 < |q| < \tfrac{1}{2}. \end{aligned} \tag{4.7}$$

Here, as in the Gaussian model, a renormalized field σ'_q depending on rescaled momenta has been defined.

There is a correlation function $\Gamma'_{q'}$ associated with the new Hamiltonian $\mathcal{H}'$; the new correlation length is,

$$\xi' = \left\{\frac{d\Gamma'}{dq^2}\frac{1}{\Gamma'}\right\}^{1/2}_{q^2=0}. \tag{4.8}$$

As in the previous section, one will have $\Gamma_q = \zeta^2 2^{-d}\Gamma'_{2q}$ and $\xi' = \frac{1}{2}\xi$.

The precise relation between $\mathcal{H}$ and $\mathcal{H}'$ can be worked out perturbatively if u is small. The most general form for $\mathcal{H}'$ can be written

$$\mathcal{H}' = -\tfrac{1}{2}\int_q u'_2(q)\,\sigma'_q\sigma'_{-q} - \int_{q_1}\int_{q_2}\int_{q_3} u'_4(q_1,q_2,q_3,-q_1-q_2-q_3)\,\sigma'_{q_1}\sigma'_{q_2}\sigma'_{q_3}\sigma'_{-q_1-q_2-q_3} + \text{higher order terms in } \sigma'. \tag{4.9}$$

The functions u'_2, u'_4, etc. will be determined from (4.7). The simplifications needed to put (4.9) into the form of (4.2) will be introduced later. Write (4.7) in the form,

$$\exp\{\mathcal{H}'[\sigma']\} = \int_{\sigma_1}\exp\{\mathcal{H}_F[\sigma] + \mathcal{H}_I[\sigma]\}$$

$$\exp\{\mathcal{H}'[\sigma']\} = \exp\{\mathcal{H}_F[\sigma_0]\}\int_{\sigma_1}\exp\{\mathcal{H}_F[\sigma_1] + \mathcal{H}_I[\sigma_0+\sigma_1]\}. \tag{4.10}$$

The term $\exp\{\mathcal{H}_F[\sigma_0]\}$ can be written trivially in terms of σ',

$$\mathcal{H}_F[\sigma_0] = -\tfrac{1}{2}\int_{0<|q|<\frac{1}{2}}(q^2+r)\,\sigma_{-q}\sigma_q \tag{4.11}$$

$$\mathcal{H}_F[\sigma_0] = -\tfrac{1}{2}(\zeta^{+2}2^{-d-2})\int_{0<|q|<1}(q^2+4r)\,\sigma'_q\sigma'_{-q}. \tag{4.12}$$

The nontrivial physics lies in the terms depending on u. To obtain these terms, consider (4.10) and expand $\exp\{\mathcal{H}_I[\sigma]\}$ in powers of u. It is convenient to write this expansion graphically. Denote $(-\mathcal{H}_I)$ by a cross (fig. 4.1a). The four endpoints of the cross symbolize the four σ's in $\mathcal{H}_I$.

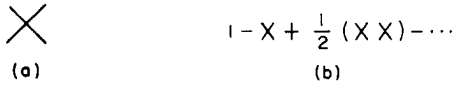

Fig. 4.1. (a) Graphical representation of $-\mathcal{H}_I$.
(b) Graphs for $\exp\{\mathcal{H}_I\}$.

The expansion

$$\exp(\mathcal{H}_I) = 1 + \mathcal{H}_I + \tfrac{1}{2}\mathcal{H}_I^2 + \tfrac{1}{6}\mathcal{H}_I^3 + \ldots \tag{4.13}$$

is shown graphically in fig. 4.1b. Consider now the functional integral of eq. (4.10). Using the expansion of (4.13), one has to calculate Gaussian integrals of the type

$$I(q_1,\ldots,q_k) = \int_{\sigma_1}\sigma_{1,q_1}\ldots\sigma_{1,q_k}\exp\{\mathcal{H}_F[\sigma_1]\}. \tag{4.14}$$

This integral can be written in terms of the functional integral

$$Z_F = \int_{\sigma_1} \exp\{\mathcal{H}_F[\sigma_1]\} \tag{4.15}$$

and the contraction $\overbrace{\sigma_{1,q}\sigma_{1,q_1}}$ which is

$$\overbrace{\sigma_{1,q}\sigma_{1,q_1}} = (2\pi)^d\delta^d(q+q_1)/(q^2+r) \tag{4.16}$$

provided that $\frac{1}{2} < |q| < 1$ (it is 0 for $|q| < \frac{1}{2}$). The integral $I(q_1, \ldots, q_k)$ is Z_F times the sum of all contractions of $\sigma_{1,q_1} \ldots \sigma_{1,q_k}$, as explained in section 3. (The derivation of section 3 for discrete spin integrations applies also to functional integrals.) The factor Z_F contributes only a constant independent of σ' to $\mathcal{H}'$ and will be dropped (we will study only σ'-dependent terms in $\mathcal{H}'$).

In the graphical expansion of fig. 4.1b each endpoint of each cross represents a variable σ_q. Writing $\sigma_q = \sigma_{0,q} + \sigma_{1,q}$, one has to consider two choices for each endpoint. When it represents $\sigma_{1,q}$, it must be contracted with another endpoint representing σ_1 also. To symbolize this one draws a line connecting the two endpoints, as in fig. 4.2b. Such a line is called a propagator. If

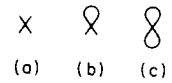

Fig. 4.2. First order graphs after performing σ_1 functional integral.

an endpoint represents $\sigma_{0,q}$ it is left as is. The functional integral in eq. (4.10) is obtained graphically by considering all possible ways of contracting some endpoints and leaving some endpoints uncontracted. Topologically equivalent graphs are then lumped together. For example the integral of $\mathcal{H}_I$ itself involves the three graphs shown in fig. 4.2. When one forms all possible contractions on the single cross, the graph with no contractions (fig. 4.2a) occurs once, the graph with one contraction (fig. 4.2b) occurs 6 times, while the graph with two contractions (fig. 4.2c) occurs 3 times. The diagrams that occur in second order are shown in fig. 4.3 together with the number of times each graph occurs (the numerical factor includes the factor $\frac{1}{2}$ multiplying $\mathcal{H}_I^2$ in eq. (4.13)).

A given diagram is calculated as follows. Consider for example the first diagram of fig. 4.3c. One starts with two crosses; before integration over σ_1 these correspond to

$$\begin{aligned} u^2 \int_q \ldots \int_{q_3} (2\pi)^d\delta^d(q+q_1+q_2+q_3)\,\sigma_q\sigma_{q_1}\sigma_{q_2}\sigma_{q_3} \\ \times \int_{q_4} \ldots \int_{q_7} (2\pi)^d\delta^d(q_4+q_5+q_6+q_7)\,\sigma_{q_4}\sigma_{q_5}\sigma_{q_6}\sigma_{q_7}. \end{aligned} \tag{4.17}$$

For each external line a σ is replaced by a σ_0, while for each internal line the corresponding σ's are replaced by a propagator. For the example the result is

$$\begin{aligned} u^2 \int_q \ldots \int_{q_7} (2\pi)^d\delta^d(q+q_1+q_2+q_3)(2\pi)^d\delta^d(q_4+\ldots+q_7)(2\pi)^d\delta^d(q_2+q_6) \\ \times (2\pi)^d\delta^d(q_3+q_7)[q_2^2+r]^{-1}[q_3^2+r]^{-1}\sigma_{0,q}\sigma_{0,q_1}\sigma_{0,q_4}\sigma_{0,q_5}. \end{aligned} \tag{4.18}$$

with $|q|$, $|q_1|$, $|q_4|$, and $|q_5|$ restricted to be less than $\frac{1}{2}$ while $|q_2|$, $|q_3|$, $|q_6|$, and $|q_7|$ must lie between $\frac{1}{2}$ and 1. Eliminating all the δ-functions, and putting in the numerical factor 36, one has

$$36 \int_q \int_{q_1} \int_{q_4} \sigma_{0,q}\sigma_{0,q_1}\sigma_{0,q_4}\sigma_{0,-q-q_1-q_4} \int_{q_2} \frac{1}{q_2^2+r} \frac{1}{(q+q_1+q_2)^2+r} \tag{4.19}$$

where q_2 is restricted so that $|q_2|$ and $|q+q_1+q_2|$ lie between $\frac{1}{2}$ and 1.

The expression (4.19) is the contribution of one diagram to the integral $\int_{\sigma_1} \exp\{\mathcal{H}_F[\sigma_1] + \mathcal{H}_I[\sigma_0+\sigma_1]\}$ of eq. (4.10). Other diagrams are calculated similarly. However what one wants to compute is $\mathcal{H}'[\sigma']$. This requires taking the logarithm of the integral and replacing σ_0 by σ'. This can be done within the diagrammatic framework. Taking the logarithm of the integral is equivalent to removing all disconnected diagrams. In second order in u, the disconnected diagrams are those of fig. 4.3a. Thus, to second order in u the logarithm of the integral involves the diagrams of fig. 4.1 plus those of figs. 4.2a–4.2c. This is easily checked explicitly. It is straightforward but tedious to verify that if $\mathcal{H}_I'$ is the sum of all connected graphs to all orders in u then $\exp\{\mathcal{H}_I'\}$ is the sum of all connected and disconnected graphs to all orders in u.

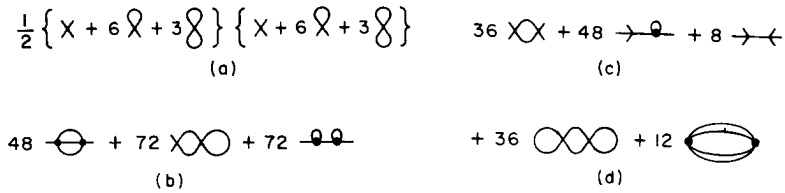

Fig. 4.3. Second order graphs after σ_1 integration.

Since we shall study only σ'-dependent terms in $\mathcal{H}'[\sigma']$ the graphs without external lines (see figs. 4.2c and 4.3d) will also be dropped (one would have to keep these graphs to calculate the free energy itself, but they are irrelevant if one computes only the spin-spin correlation function). Finally, replacing σ_0 by σ' simply means substituting $\zeta\sigma'_{2q}$ for an external line instead of $\sigma_{0,q}$. One then has to make a change of variables from q to $2q$ in order to put $\mathcal{H}'$ in the form of eq. (4.9).

One can write a set of rules for computing any diagram:

1. Label the momenta in the "incoming" sense at each vertex.
2. Internal momenta range from $\frac{1}{2}$ to 1, external momenta range from 0 to $\frac{1}{2}$.
3. Associate a propagator $(2\pi)^d \delta^{(d)}(q_1+q_2)/(q_1^2+r^2)$ with each internal line (q_1 and q_2 are the two momentum labels on that line).
4. Associate a factor $u(2\pi)^d\delta^{(d)}(q_1+q_2+q_3+q_4)$ with each four-point vertex.
5. Associate a spin variable $\sigma_{0,q} = \zeta\sigma'_{2q}$ with each external leg.
6. Integrate $\int_q$ over internal and external momenta according to rule 2.

Writing explicitly only some of the diagrams in figs. 4.2 and 4.3, the functions u_2', u_4', etc. in the new interaction $\mathcal{H}'$ (eq. (4.9)) are:

$$u_2'(q) = \zeta^{+2}2^{-d}\left\{\tfrac{1}{4}q^2 + r + 12u \int_p \frac{1}{p^2+r} - 96u^2 \int_p \int_{p_1} \frac{1}{(p^2+r)} \frac{1}{(p_1^2+r)} \frac{1}{[(\frac{1}{2}q-p-p_1)^2+r]} + \ldots\right\} \tag{4.20}$$

$$u_4'(q_1,q_2,q_3,q_4) = \zeta^{+4}2^{-3d}\left\{u - 12u^2 \int_p \frac{1}{(p^2+r)} \frac{1}{[(\frac{1}{2}q_1+\frac{1}{2}q_2-p)^2+r]} - 2 \text{ permutations } \ldots\right\} \tag{4.21}$$

$$u'_6(q_1, q_2, q_3, q_4, q_5, q_6) = -\zeta^{+6} 2^{-5d} u^2 \left\{ \frac{0.8}{[(\frac{1}{2}q_1 + \frac{1}{2}q_2 + \frac{1}{2}q_3)^2 + r]} + 9 \text{ permutations } \ldots \right\}. \quad (4.22)$$

The momenta $|p|$ and $|p_1|$ are restricted to lie between $\frac{1}{2}$ and 1. Likewise one must have

$$\tfrac{1}{2} < |\tfrac{1}{2}q - p - p_1| < 1, \qquad \tfrac{1}{2} < |\tfrac{1}{2}q_1 + \tfrac{1}{2}q_2 + \tfrac{1}{2}q_3| < 1, \qquad \text{etc.} \quad (4.23)$$

So far the discussion has been general. Now some approximations will be introduced to simplify the effective Hamiltonian to the form (4.2). These approximations will be good for dimension d near 4, but this will only become apparent later. First calculate $u'_2(q)$ only to order u. One can write

$$u'_2(q) = q^2 + r' \quad (4.24)$$

provided that

$$\zeta = 2^{1+d/2} \quad (4.25)$$

as in the preceeding section. We will further simplify (4.23) by making the approximations

$$\int_{\frac{1}{2}<|p|<1} \frac{1}{p^2 + r} \to \frac{1}{1+r} \int_{\frac{1}{2}<|p|<1} 1 = \tfrac{1}{4} c \, \frac{1}{1+r}$$

where $c = 4 \int_p 1$ with $\frac{1}{2} < |p| < 1$, and

$$\int_{\frac{1}{2}<|p|<1} \frac{1}{(p^2 + r)} \frac{1}{[(\frac{1}{2}q_1 + \frac{1}{2}q_2 - p)^2 + r]} \to \frac{1}{(1+r)^2} \int_{\frac{1}{2}<|p|<1} 1 = \tfrac{1}{4} c \, \frac{1}{(1+r)^2}.$$

If we abbreviate $u'_4(q_1 \ldots q_4)$ by u', (4.23) becomes

$$\begin{aligned} r' &= 4\left[r + 3c \, \frac{u}{1+r}\right] + \text{higher orders in } u, \\ u' &= 2^{4-d}\left[u - 9c \, \frac{u^2}{(1+r)^2}\right] + \text{higher orders in } u. \end{aligned} \quad (4.26)$$

Finally u'_6 and higher order terms will be ignored in this section. A discussion of the validity of these approximations will occupy the next section. Now eq. (4.26) will be studied as an example of the Renormalization Group at Work. (These equations were studied in [113].)

4.1. *Simplified renormalization group transformation*

The first point to understand is that eq. (4.26) is to be iterated many times. That is, if the parameters of the initial interaction are $u = u_0$ and $r = r_0$, then one uses eq. (4.26) to define a sequence of interaction parameters u_l and r_l, corresponding to a sequence of effective interactions $\mathcal{H}_l$. The effective interaction $\mathcal{H}_l$ is the result of integrating all spin components σ_q with $2^{-l} < |q| < 1$ in the original interaction. The effective interaction $\mathcal{H}_l$ describes the behavior of spin components σ_q with $0 < |q| < 2^{-l}$. Eqs. (4.26) now read

$$\begin{aligned} r_{l+1} &= 4\{r_l + 3c \, u_l/(1 + r_l)\} \\ u_{l+1} &= 2^{4-d}\{u_l - 9c \, u_l^2/(1 + r_l)^2\} \end{aligned} \quad (4.27)$$

neglecting higher orders in u_l.

4.2. *The ϵ-expansion and a non-trivial fixed point*

Consider now the fixed points of (4.26). At a fixed point ($r = r^*$, $u = u^*$), one has $r' = r$ and $u' = u$, which implies that $\xi' = \xi$. However, we have argued before that $\xi' = \frac{1}{2}\xi$, so at a fixed point the correlation length must be infinite, i.e. we must be at a critical point of the theory. (Another solution is $\xi = \xi' = 0$, but this possibility can be ruled out by other considerations.) It is clear from (4.26) that a trivial fixed point $r^* = u^* = 0$ exists as in the Gaussian model. We will see below, however, that the interactions in the present model can generate more interesting possibilities.

First imagine that the number of dimensions of the system, d, is greater than 4. If (4.27) is iterated many times u_l approaches zero, and if one approaches a fixed point it is the trivial point $r^* = u^* = 0$. So to get more exciting results, suppose $d < 4$ and u_0 is chosen small. Upon iteration u_l will increase with l until the second terms in (4.27) become comparable to the first. In this case a new, nontrivial fixed point emerges. The fixed point (u^*, r^*) is given by,

$$u^* \approx \frac{1}{9c}(2^{4-d} - 1), \qquad r^* \approx -4c\,u^*. \qquad \text{(This assumes } r^* \ll 1.\text{)} \tag{4.28}$$

If $d \approx 4$, u^* and r^* are both small and our approximations are reasonable. (See the next section.) Define

$$\epsilon = 4 - d. \tag{4.29}$$

Then, the position of the fixed point for small ϵ is

$$u^* \approx \frac{1}{9c}\,\epsilon \ln 2, \qquad r^* \approx -\tfrac{4}{9}\,\epsilon \ln 2. \tag{4.30}$$

Imagine beginning the iteration scheme (4.27) with $u_0 > u^*$. Subsequent values of u_l will decrease towards u^* (fig. 4.4a). If $u_0 < u^*$, subsequent values of u_l will increase towards u^* (fig. 4.4b). Therefore, the value of u occurring in the Hamiltonian after many iteration steps is

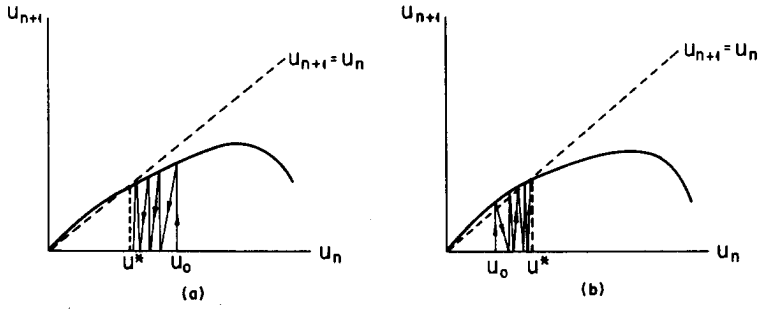

Fig. 4.4. The iteration formula for u_n. If $u_0 > u^*$, the iteration scheme (a) leads u_n to u^* from above. If $u_0 < u^*$, u_n approaches u^* from below (b).

near u^* no matter what u_0 was. This means that the fixed point value u^* determines the strength of the coupling among the long wavelength fluctuations in the system, independently of the value of u_0.

4.3. *Linearized equations and calculation of ν*

Now the calculation of ν will be discussed using eqs. (4.27). This is somewhat more complicated than in previous examples and will be worked through in detail. See also [114] and [111].

First it is useful to compare the solutions of the recursion formula (4.27) with the corresponding recursion formula of the Gaussian case $r_{l+1} = 4r_l$. In the Gaussian case one has two types of solutions. If $r_0 = 0$ then $r = 0$ for all l: This is the fixed point solution and corresponds to the case $T = T_c$. If $r_0 \neq 0$ then $r_l = 4^l r_0$ and $r_l \to \infty$ for $l \to \infty$. This corresponds to the case $T \neq T_c$. For T near T_c, r_0 is linear in $T - T_c$.

One complication of the non-Gaussian case is that one can be at the critical temperature $T = T_c$ without having r_0 and u_0 equal to the fixed point values. The reason is that one has to fix only one parameter T to be at the critical temperature, while one must fix both $r_0 = r^*$ and $u_0 = u^*$ to be at the fixed point. Thus one can imagine that for *any* value of u_0, there will be a critical value r_{0c} for r_0 (r_{0c} will depend on u_0) which corresponds to $T = T_c$. The role of the fixed point for $T = T_c$ is that $r_l \to r^*$ and $u_l \to u^*$ as $l \to \infty$. (If $T \neq T_c$, r_l and u_l will have another limiting behavior for large l.)

Suppose one is considering a specific starting interaction with T the only variable. Then $r_0 = r_0(T)$ and $u_0(T)$ will depend on T in a specific analytic way. (See, e.g. (3.29).) Write $r_0(T_c) = r_{0c}$, $u_0(T_c) = u_{0c}$. We want now to study the theory for T near T_c, in order to calculate the exponent ν. The first step is to study the sequence $\{r_l(T), u_l(T)\}$ generated by iterating (4.27) starting from $r_0(T)$, $u_0(T)$. Examples of these sequences are shown in fig. 4.5. We expect $r_l(T_c) \to r^*$, $u_l(T_c) \to u^*$ as $l \to \infty$ (line A of fig. 4.5). Since the recursion equations are analytic, one also expects $r_l(T) = r_l(T_c) + (T - T_c) r_l'(T_c)$ (and similarly, for $u_l(T)$) for fixed l and small enough $T - T_c$. If l is sufficiently large $r_l(T_c) \approx r^*$ so $r_l(T)$ is also close to r^*.

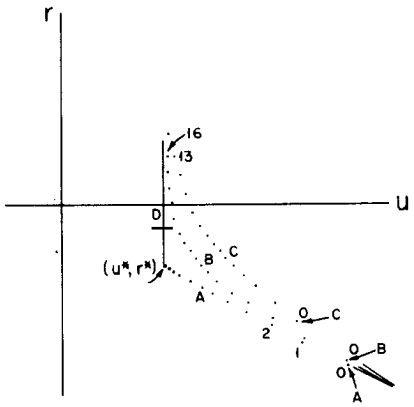

Fig. 4.5. Plot of the iteration scheme for three different initial choices of parameters. Sequence A($T = T_c$) goes into the fixed point (u^*, r^*). Sequences B and C begin for choices of u and r slightly removed from criticality. These sequences eventually deviate far from the fixed point but approach the unique curve D.

This suggests that we study the recursion formulae for the case $r_l \approx r^*$, $u_l \approx u^*$. In this case one can linearize the recursion formulae. The result, in matrix form, is

$$\begin{bmatrix} r_{l+1} - r^* \\ u_{l+1} - u^* \end{bmatrix} \approx M \begin{bmatrix} r_l - r^* \\ u_l - u^* \end{bmatrix} \tag{4.31}$$

where

$$M = \begin{bmatrix} 4 - \dfrac{12cu^*}{(1+r^*)^2} & \dfrac{12c}{1+r^*} \\ +\dfrac{2^\epsilon \cdot 18cu^{*2}}{(1+r^*)^3} & 2^\epsilon\left[1 - \dfrac{18cu^*}{(1+r^*)^2}\right] \end{bmatrix}. \tag{4.32}$$

It will be useful to know the result of iterating the linearized recursion relation many times. The result is, formally

$$\begin{bmatrix} r_{l+n} - r^* \\ u_{l+n} - r^* \end{bmatrix} = M^n \begin{bmatrix} r_l - r^* \\ u_l - u^* \end{bmatrix}. \tag{4.33}$$

The advantage of studying M^n, with n large, instead of M itself is that M^n is completely dominated by the largest eigenvalue of M. The eigenvalues of M are easily calculated. In the limit $u^* \to 0$, the eigenvalues are 4 and 1; hence the eigenvalues of M^n are 4^n and 1; clearly the eigenvalue 4^n dominates.

We need the explicit form of M^n neglecting the eigenvalue 1. This is easily obtained once M has been diagonalized. The diagonalization is complicated because M is not symmetric; the result can be written as follows

$$M_{ij} = \lambda_1 w_{i1} v_{1j} + \lambda_2 w_{i2} v_{2j} \tag{4.34}$$

where w_{il} and v_{li} are eigenvectors (with eigenvalue λ_l) of the matrices M and M^{T}, respectively. To be specific,

$$\sum_j M_{ij} w_{jk} = \lambda_k w_{ik}, \qquad \sum_i v_{ki} M_{ij} = \lambda_k v_{kj}, \tag{4.35}$$

and the orthonormality relation is

$$\sum_j v_{kj} w_{jl} = \delta_{kl}. \tag{4.36}$$

Not every asymmetric matrix has a complete set of such eigenvectors, but for the matrix M the eigenvectors can be constructed explicitly. (See table 4.1.) In terms of w and v and λ, one has

$$(M^n)_{ij} = \sum_k \lambda_k^n w_{ik} v_{kj}. \tag{4.37}$$

Table 4.1
Eigenvalues and eigenvectors of the matrix M (to order ϵ)

$\lambda_1 = 4 - \frac{8}{3}\epsilon \ln 2$ $\qquad$ $\lambda_2 = 1 - \epsilon \ln 2$

$w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\qquad$ $v_1 = \begin{bmatrix} 1 \\ 4c(1 + \frac{5}{9}\epsilon \ln 2) \end{bmatrix}$

$w_2 = \begin{bmatrix} -4c(1 + \frac{5}{9}\epsilon \ln 2) \\ 1 \end{bmatrix}$ $\qquad$ $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

If λ_1 is the eigenvalue near 4 (for small u^*) while λ_2 is the eigenvalue near 1, then for large n,

$$\begin{aligned} r_{l+n} - r^* &\approx \lambda_1^n w_{11}[v_{11}(r_l - r^*) + v_{12}(u_l - u^*)] \\ u_{l+n} - u^* &\approx \lambda_1^n w_{21}[v_{11}(r_l - r^*) + v_{12}(u_l - u^*)]. \end{aligned} \tag{4.38}$$

What is the significance of this result? Let l be large but fixed, let T be near enough to T_c so that $r_l - r^*$ and $u_l - u^*$ are small and it is legitimate to use the linearized recursion equation. First let $T = T_c$. Then it was stated that $r_l \to r^*$ and $u_l \to u^*$ for $l \to \infty$. This is consistent with eq. (4.38) only if

$$v_{11}[r_l(T_c) - r^*] + v_{12}[u_l(T_c) - u^*] = 0. \tag{4.39}$$

Furthermore λ_2 must be <1, otherwise the λ_2^n term in M would grow with n also. One can see from table 4.1 that $\lambda_2 < 1$ (for $\epsilon > 0$).

Now consider the case $T \approx T_c$. For l fixed, $r_l(T) - r_l(T_c)$ and $u_l(T) - u_l(T_c)$ will both be proportional to $T - T_c$ so that,

$$v_{11}[r_l(T) - r^*] + v_{12}[u_l(T) - u^*] = c_l(T - T_c) \tag{4.40}$$

where c_l is a constant. Hence, from (4.38),

$$r_{l+n} - r^* = \lambda_1^n w_{11} c_l (T - T_c), \qquad u_{l+n} - u^* = \lambda_1^n w_{21} c_l (T - T_c). \tag{4.41}$$

This equation is valid provided $r_{l+n} - r^*$ and $u_{l+n} - u^*$ are small; otherwise one could not have used the linearized recursion formula to compute r_{l+n} and u_{l+n}.

The implications of (4.41) are illustrated graphically in fig. 4.5. The curves B and C in fig. 4.5 represent two possible trajectories (r_l, u_l). For large l they approach asymptotically the unique curve D; eq. (4.41) is the result of replacing the asymptotic part of B and C by D. The only remaining distinction between B and C is where a given point (r_l, u_l) is located on D; this is determined by the factor $c_l(T - T_c)$.

Now one can calculate ν [111]. This is done as follows. The correlation length ξ is defined for any choice of the interaction parameters r and u; write,

$$\xi = X(r, u). \tag{4.42}$$

In particular, it is defined for $r = r_0(T)$, $u = u_0(T)$ giving a function

$$\xi = \xi_0(T), \tag{4.43}$$

say. Now the scaling rule for effective interactions gives

$$X(r_{l+n}, u_{l+n}) = 2^{-l-n} X(r_0, u_0) \tag{4.44}$$

if r_{l+n} and u_{l+n} are solutions of the recursion formula. But from eq. (4.41) one sees the following: $r_{l+n+1}(T) - r^*$, for $T - T_c = \tau/\lambda_1$, is the same as $r_{l+n}(T) - r^*$ for $T - T_c = \tau$. Likewise for u_{l+n+1}. Thus

$$X[r_{l+n+1}, u_{l+n+1}]\big|_{T = T_c + \tau/\lambda_1} = X[r_{l+n}, u_{l+n}]\big|_{T = T_c + \tau}. \tag{4.45}$$

This means that

$$2^{-l-n-1}\xi(T_c + \tau/\lambda_1) = 2^{-l-n}\xi(T_c + \tau). \tag{4.46}$$

Now if

$$\xi(T_c + \tau) \propto \tau^{-\nu} \tag{4.47}$$

one has for arbitrarily small τ:

$$(\tau/\lambda_1)^{-\nu} = 2(\tau^{-\nu}). \tag{4.48}$$

This equation can be true only if

$$\lambda_1^{\nu} = 2, \tag{4.49}$$

i.e.,

$$\nu = \ln 2/\ln \lambda_1. \tag{4.50}$$

To order ϵ this gives

$$\nu = \ln 2/(2 \ln 2 - \tfrac{1}{3}\epsilon \ln 2) = 0.5 + \tfrac{1}{12}\epsilon + O(\epsilon^2).$$

A few remarks about this result are in order. The artificial features of the iteration scheme (namely, the factors of log 2) do not appear in ν. This is crucial since ν should be property of the physical system and not of the method of solution! Also, for non-zero ϵ, ν differs from the Gaussian (mean field) model. Hence, the model studied in this section is an example in which interactions play a role in determining the critical behavior of a physical system. The value of ν is determined through (4.50) in terms of the largest eigenvalue λ_1 of the matrix M. This is similar to the results of section 2. In section 2 the 2 x 2 matrix M is replaced by a 1 x 1 matrix with a single eigenvalue λ (see (2.26)). Analogously, the equation of section 2 for ν (2.29) is obtained from (4.50) by substituting λ for λ_1. Finally, note that ν is independent of the initial constant u_0. This is an example of the universality hypothesis discussed in section 1.

5. The s^4 model (cont'd.)

In the preceding section the renormalization group equations were mutilated in order to arrive at results quickly. In this section the renormalization group equations will be studied more carefully. The end result will be that the simple equations of the preceding section are correct to order $\epsilon = 4 - d$. The discussion will be in two parts. First, abstract generalizations of the simplified equations will be examined to find out which generalizations could cause trouble. Then the correct renormalization group equations will be examined.

5.1. *Irrelevant variables and the ϵ expansion*

The recursion formulae of the preceding section were,

$$r' = 4[r + 3cu/(1 + r)], \qquad u' = 2^{\epsilon}[u - 9cu^2/(1 + r)^2] \tag{5.1}$$

where $\epsilon = 4 - d$ and d is the dimensionality of the system. These equations possess a fixed point,

$$u^* = \frac{1}{9c}\epsilon \ln 2, \qquad r^* = -\tfrac{4}{9}\epsilon \ln 2 \quad \text{(to order } \epsilon\text{)}. \tag{5.2}$$

Suppose that the true equations had one other coupling constant w. There will then be another equation for w'. Suppose that (at least near the fixed point) the equation for w' is,

$$w' = aw \tag{5.3}$$

where $a < 1$. Then, when (5.3) is iterated, $w' \to 0$ and the variable w disappears from the problem. w is then referred to as an *irrelevant variable* with respect to the fixed point for which (5.3) is true. The term "irrelevant variable" is not to be taken literally; the meaning of "irrelevant" in critical phenomena is complicated, but should be clarified by this and subsequent sections. For another discussion of irrelevant variables and fixed points, see [114]. The term "irrelevant variable" is due to Kadanoff [115].

Now we want to know when the presence of irrelevant variables can change the fixed point (5.2). Suppose that the equation for w' is,

$$w' = \tfrac{1}{4} w + (\text{second order terms like } u^2, uw, w^2, \text{etc.}). \tag{5.4}$$

In particular, we forbid the appearance of a term linear in u in (5.4). (w is still classified as an irrelevant variable: second order terms are to be ignored in determining whether w is relevant or irrelevant. A term proportional to u can only appear in the formula for r' or u' since the only diagrams linear in u are figs. 5.1a, b.)

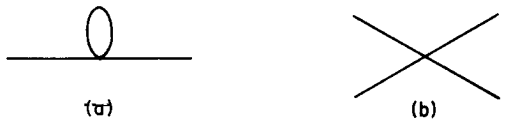

Fig. 5.1. Diagrams linear in u.

Consider the fixed point of (5.4). Obviously,

$$w^* = O(\epsilon^2) \tag{5.5}$$

assuming that (5.2) does not fail so that $u^* = O(\epsilon)$. We must now ask whether (5.5) can invalidate the eq. (5.2) for u^* and r^* if terms involving w should appear in (5.1). Only terms of order ϵ are important in the equation for r^*, so w^* is negligible. However, the u equation involves ϵ^2 terms in an important way. Therefore, if w appeared *linearly* in the u equation, (5.2) would be modified in an important way.

Now consider the case that the w' equation has the form,

$$w' = \tfrac{1}{4} w + \text{terms of order } u^3, u^2 w, \text{ etc.} \tag{5.6}$$

In this case $w^* \sim O(\epsilon^3)$ and even if w appeared linearly in the u equation, it could not influence the validity of that equation for small ϵ.

These considerations show that the simple equation for u' can be invalidated if there is an irrelevant variable w such that both (1) $w^* \sim O(\epsilon^2)$, and (2) w enters the u' equation linearly. We will now argue that this possibility does not occur for the exact equations defined in the previous section. In the exact equations, the interaction (after some iterations) has the form

$$\mathcal{H} = -\tfrac{1}{2} \int_q u_2(q)\sigma_q \sigma_{-q} - \int_{q_1} \int_{q_2} \int_{q_3} \int_{q_4} \delta(q_1 + q_2 + q_3 + q_4) u_4(q_1, q_2, q_3, q_4) \sigma_{q_1} \sigma_{q_2} \sigma_{q_3} \sigma_{q_4} - (\text{6th order term}) - (\text{8th order term}) - \ldots . \tag{5.7}$$

As discussed at length in the previous several sections, we integrate out the high frequency modes in this system. After the appropriate rescaling of momenta and fields, a primed system emerges which is described by a Hamiltonian identical in form with (5.7), but expressed in terms of primed quantities,

$$\mathcal{H}' = -\tfrac{1}{2}\int_q u_2'(q)\sigma_q'\sigma_{-q}' - \int_{q_1}\int_{q_2}\int_{q_3}\int_{q_4}\delta(q_1+q_2+q_3+q_4)u_4'(q_1,q_2,q_3,q_4)\sigma_{q_1}'\sigma_{q_2}'\sigma_{q_3}'\sigma_{q_4}' - (\text{6th order term})' - \ldots . \tag{5.8}$$

The relations between the u_i' and u_i follow by the techniques illustrated in section 4 (assuming $u_2(\boldsymbol{p})$ depends only on $|\boldsymbol{p}| = p$ because of the rotational symmetry of these equations):

$$u_2'(q) = \zeta^2 2^{-d}\left\{u_2(\tfrac{1}{2}q) + 12\int_p \frac{u_4(\frac{1}{2}q, -\frac{1}{2}q, p, -p)}{u_2(p)} + (\text{second order term}) + \ldots\right\},$$

$$u_4'(q_1,q_2,q_3,q_4) = \zeta^4 2^{-3d}\left\{u_4(\tfrac{1}{2}q_1,\tfrac{1}{2}q_2,\tfrac{1}{2}q_3,\tfrac{1}{2}q_4)\right.$$
$$-12\int_p \frac{u_4(\frac{1}{2}q_1,\frac{1}{2}q_2,p,-\frac{1}{2}q_1-\frac{1}{2}q_2-p)u_4(\frac{1}{2}q_3,\frac{1}{2}q_4,-p,-\frac{1}{2}q_3-\frac{1}{2}q_4+p)}{u_2(p)u_2(\frac{1}{2}q_1+\frac{1}{2}q_2+p)} + \text{perms.}$$
$$\left.+15\int_p \frac{u_6(\frac{1}{2}q_1,\frac{1}{2}q_2,\frac{1}{2}q_3,\frac{1}{2}q_4,p,-p)}{u_2(p)} + \ldots\right\}$$

$$u_6'(q_1,\ldots,q_6) = \zeta^6 2^{-5d}\left\{u_6(\tfrac{1}{2}q_1,\ldots,\tfrac{1}{2}q_6)\right. \tag{5.9}$$
$$\left.+0.8\,\frac{u_4(\frac{1}{2}q_1,\frac{1}{2}q_2,\frac{1}{2}q_3,-\frac{1}{2}q_1-\frac{1}{2}q_2-\frac{1}{2}q_3)u_4(\frac{1}{2}q_4,\frac{1}{2}q_5,\frac{1}{2}q_6,-\frac{1}{2}q_4-\frac{1}{2}q_5-\frac{1}{2}q_6)}{u_2(\frac{1}{2}q_1+\frac{1}{2}q_2+\frac{1}{2}q_3)} + \ldots\right\}.$$

Whenever u_2 appears in a denominator, its argument must lie between $\frac{1}{2}$ and 1, e.g. $\frac{1}{2} < p < 1$, etc. Clearly the system of equations (5.9) is intractable. The equation for u_4' in particular is a non-linear functional equation. Therefore, the contention that (5.9) can be replaced by algebraic equations (5.1) (to lowest order in ϵ) is extremely important. The proof that the algebraic equations are sufficient to determine the fixed point r^*, u^*, rests heavily on the character of the irrelevant variable u_6.

An exact fixed point of (5.9) is a set of fixed functions $u_2^*(q)$, $u_4^*(q_1,\ldots,q_4)$, $u_6^*(q_1,\ldots,q_6)$, etc. To test whether the simple equations are a valid approximation, we will suppose that to order ϵ these fixed functions reduce to two fixed constants r^* and u^*, namely,

$$u_2^*(q) = r^* + q^2 + O(\epsilon^2)$$
$$u_4^*(q_1,\ldots,q_4) = u^* + O(\epsilon^2) \tag{5.10}$$
$$u_6^*(q_1,\ldots,q_6) = O(\epsilon^2)$$

etc. The consistency of this assumption will be verified below.

First consider the u_2 equation. Write

$$u_2^*(q) = q^2 + r^* + v_2^*(q) \tag{5.11}$$

where $v_2^*(q)$ is a remainder. The parameter ζ will be adjusted to ensure that the coefficient of q^2 is unity. Let (5.10) represent an expansion about $q^2 = 0$, so that

$$v_2^*(q) \sim q^4 \cdot O(\epsilon^2). \tag{5.12}$$

Substituting (5.11) into the equation for $u_2(q)$ gives,

$$q^2 + r^* + v_2^*(q) = \zeta^2 2^{-d}\left\{\tfrac{1}{4}q^2 + r^* + v_2^*(\tfrac{1}{2}q) + 12 \int_p \frac{u_4^*(\frac{1}{2}q, -\frac{1}{2}q, p, -p)}{u_2^*(p)} + \ldots\right\}. \tag{5.13}$$

Now, to order ϵ the ... is negligible and the last term reduces to

$$12\, u^* \int_p \frac{1}{p^2 + r^*} + O(\epsilon^2). \tag{5.14}$$

The parameter ζ is to be determined (as in previous sections) by the requirement that terms proportional to q^2 in (5.12) match. This gives

$$\zeta^2 2^{-d} = 4 + O(\epsilon^2). \tag{5.15}$$

(There is more discussion of ζ at the end of this section.) Now one has an equation for r^*,

$$r^* = 4\left\{r^* + 12\, u^* \int_p \frac{1}{p^2 + r^*} + O(\epsilon^2)\right\} \tag{5.16}$$

and an equation for the remainder $v_2^*(q)$,

$$v_2^*(q) = 4v_2^*(\tfrac{1}{2}q) + \epsilon^2 F_2^*(q). \tag{5.17}$$

In this equation $\epsilon^2 F_2^*$ simply represents the corrections from higher order terms in (5.13) (with constant and q^2 terms subtracted). Suppose that F_2^* is known; a series solution can be written for $v_2^*(q)$ in terms of $F_2^*(q)$:

$$v_2^*(q) = \epsilon^2 \sum_{n=0}^{\infty} 2^{2n} F_2^*(2^{-n}q). \tag{5.18}$$

This is a solution only if the series converges. However, F_2^* was constructed such that $F_2^*(q) \sim q^4$. Therefore, the nth term in (5.18) is of order,

$$\epsilon^2 2^{2n}\, q^4/2^{4n} = 2^{-2n}\epsilon^2 q^4 \tag{5.19}$$

and the series converges. Therefore $v_2^* \sim O(\epsilon^2)$ as supposed.

What is the connection of this discussion with irrelevant variables? To see this, expand u_2 in powers of q,

$$u_2(q) = r + q^2 + w_1 q^4 + w_2 q^6 + \ldots\,. \tag{5.20}$$

Then, (5.17) gives to lowest order

$$w_1' = \tfrac{4}{16} w_1, \qquad w_2' = \tfrac{4}{64} w_2 \tag{5.21}$$

and *all* the variables $w_1, w_2, \ldots$ are irrelevant variables.

It should be clear now that v_2^* cannot upset the simple equations for r^* and u^*, since it appears only in the terms with propagators in (5.9). These terms are already $O(\epsilon^2)$, so the presence of v_2^* leads to corrections $\sim O(\epsilon^4)$ which can be ignored.

The discussion of u_4 parallels u_2. We begin by writing,

$$u_4^*(q_1, \ldots, q_4) = u^* + v_4^*(q_1, \ldots, q_4) \tag{5.22}$$

where v_4^* is a remainder defined such that $v_4^*(0 \ldots 0) = 0$. Then, substituting (5.22) into (5.9) and using (5.15), we can obtain equations for u^* and $v_4^*(q_1, \ldots, q_4)$,

$$u^* = 2^\epsilon \left\{ u^* - 36 \int_p \frac{u^{*2}}{(p^2 + r^*)^2} + 15 \int_p \frac{u_6^*(0, 0, 0, 0, p, -p)}{p^2 + r^*} + \ldots \right\}, \tag{5.23}$$

$$v_4^*(q_1, \ldots, q_4) = 2^\epsilon v_4^*(\tfrac{1}{2}q_1, \ldots, \tfrac{1}{2}q_4) + \epsilon^2 F_4^*(q_1, \ldots, q_4). \tag{5.24}$$

The . . . terms in the equation for u^* are all of order ϵ^3 or higher and can be ignored. Eq. (5.24) implies that $v_4^* \sim O(\epsilon^2)$ in the same way that (5.17) gave $v_2^* \sim O(\epsilon^2)$. Since v_4^* does not appear linearly in the equation for u^*, v_4^* cannot invalidate the simple equations.

Now consider u_6. Since u_6 appears *linearly* in the u^* equation, it is potentially dangerous. Consider the equation (5.9) for u_6' and the term corresponding to the diagram in fig. 5.2 in particular. This term appears only when the internal line carries momentum greater than $\frac{1}{2}$, i.e.

$$|\tfrac{1}{2}q_1 + \tfrac{1}{2}q_2 + \tfrac{1}{2}q_3| > \tfrac{1}{2}. \tag{5.25}$$

However, it is only $u_6(0, 0, 0, 0, p, -p)$ which appears in the u^* equation. This corresponds to $q_1 = q_2 = 0, q_3 = p$ in fig. 5.2. In order to satisfy (5.25) $|p|$ must be precisely 1 (see fig. 5.3). This is only one point in the integral over $|p|$ in (5.23) and therefore does not contribute to u^*. There

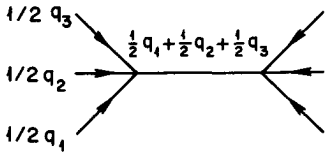

Fig. 5.2. Diagram contributing to $u_6(q_1, \ldots, q_6)$.

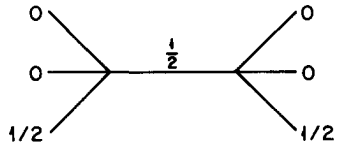

Fig. 5.3. The only diagram contributing to u_6 and entering the u^* equation linearly.

are other graphs which contribute to $u_6(0, 0, 0, 0, p, p)$ for $|p| < 1$ but these are of order ϵ^3; thus u_6 does *not* contribute an ϵ^2 term to the u^* equation.† Therefore, the u^* equation is safe!

Finally, we must cast the u^* equation into the form we had last section. In particular, since $r^* \sim O(\epsilon)$ and $u^* \sim O(\epsilon)$, we can replace,

$$\int_p \frac{u^{*2}}{(p^2 + r^*)^2} = u^{*2} \int_p \frac{1}{p^4} + O(\epsilon^3) \tag{5.26}$$

and if we redefine the constant c of section 4 to be

$$c = 4 \int_p \frac{1}{p^4} \tag{5.27}$$

we find that (5.23) reads,

$$u^* = 2^\epsilon \{u^* - 9cu^{*2}\} \tag{5.28}$$

which has the same form as (5.1) if we likewise neglect the r in the denominator of the second

†The equation for u_6 has a term linear in u_8, so if u_8 were of order ϵ^2 then u_6 would be also. But explicit study of the equation for u_8 shows that u_8 is of order ϵ^3.

term of (5.1). Now, we turn to the equation for r^*. In particular, to calculate ν one needs (5.16) for small departures from the critical point. One gets:

$$r' - r^* = 4\left\{r - r^* - 12u^* \int_p \frac{1}{(p^2 + r^{*2})}(r - r^*) + \ldots\right\}. \tag{5.29}$$

We may approximate,

$$\int_p \frac{1}{(p^2 + r^*)^2} = \int_p \frac{1}{p^4} + O(\epsilon) = \tfrac{1}{4}c + O(\epsilon) \tag{5.30}$$

and rewrite (5.29),

$$r' - r^* = 4[1 - 3u^*c](r - r^*). \tag{5.31}$$

Since the same constant c appears in both (5.28) and (5.31), it will cancel out in the calculation of the critical exponent ν as in section 4. Therefore, we have shown that (5.1) suffices to lowest order in ϵ in determining the critical exponent.

5.2. *Complete renormalization group transformation*

What one sees from this discussion is that the exact renormalization group equations are not hopelessly intractable functional equations. To order ϵ they reduce to simple algebraic equations for two parameters r^* and u^*. The basic reason that made this simplification possible was the "iterability" of the equations for the remaining functional variables $v_2^*(q)$, $v_4^*(\boldsymbol{q}, \ldots, \boldsymbol{q}_3)$, etc. That is, if the correction terms $F_2^*(q)$ in the equation for $v_2^*(q)$ are known, and are of order ϵ^2, then $v_2^*(q)$ can be computed explicitly from (5.18) and is also of order ϵ^2. This is in contrast to the equation for u^*, in this case the second order terms are of order ϵ^2 but the solution u^* is of order ϵ.

The iterability of the equations for the variables $v_2^*(q)$, etc. ensures that one can consistently calculate the exact fixed point $u_2^*(q)$, $u_4^*(\boldsymbol{q}, \ldots, \boldsymbol{q}_3)$, etc. to higher orders in ϵ. Then the functions $v_2^*(q)$, etc. cannot be neglected; but no new problems arise in calculating them (except possibly some difficulties with integrations in non-integral dimension d). This is because one can arrange a calculation so that the diagrams contributing to $v_2^*(q)$, etc. (i.e. the function $F_2^*(q)$) are known to any given order ϵ^n before $v_2^*(q)$ itself is calculated to order ϵ^n. The iterability ensures then that $v_2^*(q)$ exists in order ϵ^n.

The iterability of the equations for $v_2^*(q)$, etc. is a consequence of the fact that these functions can be expressed entirely in terms of irrelevant variables. For example, we showed that if $v_2^*(q)$ is written $w_1^* q^4 + w_2^* q^6 + \ldots$ then the equations for w_1^*, etc. are of the form

$$\begin{aligned} w_1^* &= a_1 w_1^* + \text{second order term}, \\ w_2^* &= a_2 w_2^* + \text{second order term}, \end{aligned} \tag{5.32}$$

etc. where $a_1 < 1$, $a_2 < 1$, etc. The fact that $a_i < 1$ makes it possible to solve these equations for w_1^* and w_2^* if the second order terms are known.

Iterability is possible even for relevant variables: in particular, the equation for r^* is also iterable if one writes it

$$r^* = 4r^* + \text{other terms}. \tag{5.33}$$

Then one can also solve for r^*. It is only equations like the equation for u^*:

$$u^* = a_3 u^* + \text{other terms} \tag{5.34}$$

with $a_3 = 1 + \epsilon$ which is not iterable, because a_3 is too close to 1.

Finally, consider further the constant ζ. In all calculations so far ζ^2 has turned out to be 2^{d+2} independent of ϵ. But in a higher order calculation ζ will change. The equation which determines ζ is,

$$q^2 = 2^{-d}\zeta^2\{\tfrac{1}{4}q^2 + (q^2 \text{ terms from all diagrams contributing to } u_2)\}. \tag{5.35}$$

The lowest order diagram which has q^2 dependence is shown in fig. 5.4. This diagram is of order

Fig. 5.4. Lowest order diagram which affects the value of ζ.

ϵ^2, so ζ will have to have a compensating term in order ϵ^2. It will be shown in section 7 that the value ζ^* of ζ at the fixed point determines the critical exponent η. The formula is

$$\eta = d + 2 - 2 \ln \zeta^*/\ln 2. \tag{5.36}$$

6. The approximate recursion formula

This section is devoted to the approximate recursion formula [111]. The approximate recursion formula is an approximation to the renormalization group transformation discussed in sections 4 and 5. It will be used to provide order-of-magnitude understanding of the renormalization group for $2 < d \leqslant 4$ and will shed light on the validity of the ϵ expansion.

The extra degrees of freedom in the exact renormalization group equations, as compared to the simple equations for r and u, are of two types. First of all, there are momentum dependent functions like $u_2(q)$ and $u_4(q \ldots q_3)$ instead of discrete parameters r and u. Secondly, there are an infinite set of functions u_2, u_4, u_6, u_8, etc., corresponding to all powers of the spin variable, not just u_2 and u_4. The approximate recursion formula takes into account all powers of the spin, but does not allow momentum dependence except in u_2. Most importantly for this report, the approximate recursion formula includes Feynman graphs of arbitrarily high orders and in sufficient number so that the divergent nature of the Feynman graph expansion is preserved by the recursion formula. This means the recursion formula can be used as a model to see how badly the higher order graphs affect the ϵ expansion when calculated, say, in 3 dimensions.

The approximate recursion formula was first obtained by functional integral techniques [111]. Here a more transparent graphical derivation will be given due to Polyakov [116].

6.1. *Polyakov's derivation*

Consider an interaction of the form

$$\mathcal{H} = -\tfrac{1}{2} \int_x \{\bar{\nabla} s(x)\}^2 - c_1 \int_x Q[c_2 s(x)] \tag{6.1}$$

where $s(x)$ is our usual spin field and the function Q is defined through its expansion,

$$Q[y] \equiv ry^2 + uy^4 + wy^6 + \ldots . \tag{6.2}$$

Two constants c_1 and c_2 have been introduced in $\mathcal{H}$ and will be adjusted so that the recursion formula will be free of certain phase space integrals. Note that the only momentum dependence in $\mathcal{H}$ appears in its first term, i.e. no gradients are allowed in the expression (6.2). As usual, we are going to force the renormalization group equations to produce an effective Hamiltonian $\mathcal{H}'$ with the same form as $\mathcal{H}$. That is, with the help of some approximations one tries to produce an effective Hamiltonian $\mathcal{H}'$ of the form

$$\mathcal{H}' = -\tfrac{1}{2} \int_x \{\nabla s'(x)\}^2 - c_1 \int_x Q'[c_2 s'(x)] \tag{6.1'}$$

(c_1 and c_2 are the same constants as appear in $\mathcal{H}$) where $s'(x)$ is the effective spin field obtained by Fourier transforming $\sigma'_{q'}$; the field $\sigma'_{q'}$ is as in previous lectures

$$\sigma'_{q'} = \zeta^{-1} \sigma_{q'/2}$$

and $\mathcal{H}'$ is defined by integrating out the components σ_q with $\frac{1}{2} < |q| < 1$. The only difference between the approximate recursion formula to be derived here and either the exact or simple equations of sections 4 and 5 is that different approximations will be made.

To illustrate the approximations to be made consider the graph in fig. 6.1 which contributes to u'. The prescription for neglecting momentum dependence will be to arbitrarily set the

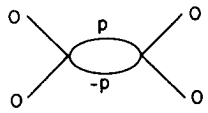

Fig. 6.1. A graph contributing to to u' in the approximate recursion formula.

momentum of the external lines to zero. This is not an absurd idea because the external lines have lower momenta ($|q| < \frac{1}{2}$) than the internal lines ($|q| > \frac{1}{2}$). So, the approximation goes in the right direction anyway. Now, fig. 6.1 gives the following contribution to u':

$$36 \left\{ \int_p u^2 (c_1 c_2^4)^2 \frac{1}{(p^2 + 2rc_1c_2^2)^2} \right\} 2^{4-d} \frac{1}{-c_1 c_2^4} . \tag{6.3}$$

The factors of $c_1 c_2^4$ multiplying u in (6.3) appear because they appear multiplying u in the interaction (6.1): to fourth order in $s(x)$, $\mathcal{H}$ is

$$= -\tfrac{1}{2} \int_x [\nabla s(x)]^2 - rc_1c_2^2 \int_x s^2(x) - uc_1c_2^4 \int_x s^4(x) + \ldots . \tag{6.4}$$

The factor $2^{4-d} = \zeta^4 2^{-3d}$ appears in (6.3) as a result of the definition of σ'_q as discussed in section 4 and the value of ζ (4.25); and finally, the last factor $(-c_1c_2^4)^{-1}$ appears in (6.3) because the strength of the $s'(x)^4$ term in $\mathcal{H}'$ is $-u'c_1c_2^4$, so to obtain u' one must divide by $-c_1c_2^4$. In

developing the approximate recursion formula contributions such as (6.3) are simplified further by applying the following two rules:

1. Replace p^2 in the integrand by an average value p_0^2.
2. Replace the integral $\int_p$ ($\frac{1}{2} < |p| < 1$) by a constant P.

Clearly, such simplifications are not necessary in discussing relatively simple graphs such as fig. 6.1. However, the neglect of the precise momentum dependence of the internal lines will allow one to calculate graphs of arbitrarily high order.

Yet another calculational rule is imposed,

3. Only even numbers of internal lines are allowed at each vertex.

The primary purpose of this rule is to exclude graphs of the form shown in fig. 6.2. Such graphs should not appear in the physical theory since they do not conserve momentum. However, rules 1 and 2 by themselves would permit such effects; this will be shown below. Rule 3 also eliminates graphs such as fig. 6.3. In the limit that the external momenta go to zero such graphs are forbidden

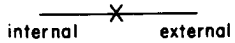

Fig. 6.2. A graph which does not appear in the approximate recursion formula. It is omitted in accordance with rule 3.

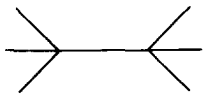

Fig. 6.3. A graph eliminated by rule 3.

by momentum conservation so this is also good. Unfortunately, this rule also forbids graphs, such as fig. 6.4, which should be present in the real theory. In particular, the class of graphs which are allowed to contribute to r' is very restrictive in this approach. It is not true, however, that all the effects of graphs of the type in fig. 6.4 will be lost. For example, after one iteration the graph in fig. 6.1 contributes to u'. In the next iteration two of the external legs of fig. 6.1 might be contracted forming a graph topologically like that in fig. 6.4. In other words, within the framework of these rules the graph of fig. 6.5 would be omitted while the graph of fig. 6.6 would appear.

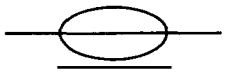

Fig. 6.4. A desirable graph forbidden by rule 3.

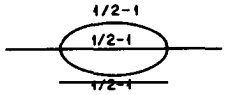

Fig. 6.5. This graph is not accounted for in the approximate recursion formula. The labels 1/2−1 indicate the momentum ranges on the internal legs.

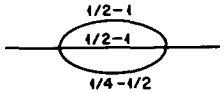

Fig. 6.6. This graph occurs in the approximate recursion formula.

Now we shall simply write down the approximate recursion relation and show from its perturbation expansion that it is equivalent to the rules 1, 2, and 3 above. The approximate recursion formula is,

$$\exp[-Q'(x)/2^d] = \frac{\int_{-\infty}^{\infty} dy \exp[-y^2 - \frac{1}{2}Q(2^{1-d/2}x + y) - \frac{1}{2}Q(2^{1-d/2}x - y)]}{\int_{-\infty}^{\infty} dy \exp[-y^2 - \frac{1}{2}Q(y) - \frac{1}{2}Q(-y)]}. \tag{6.5}$$

To compare this formula with the rules for simplifying graphs, this formula must also be expanded in diagrams. This means writing out the exponential of the numerator as

$$\exp\{-y^2 - \tfrac{1}{2}Q(y + 2^{1-d/2}x) - \tfrac{1}{2}Q(-y + 2^{1-d/2}x)\} =$$
$$\exp\{-(1+r)y^2 - 2^{2-d}rx^2 - uy^4 - 6 \times 2^{2-d}uy^2x^2 - 2^{4-2d}ux^4 - \ldots\}. \tag{6.6}$$

One then expands in powers of the quartic terms (proportional to u) and higher. This can be done by graphs, just as was done for the exact functional integral. The uy^4, $6uy^2x^2 2^{2-d}$ and $ux^4 2^{4-2d}$ terms, for example, define the 4-point vertex, with uy^4 corresponding to all lines being internal, the $6uy^2x^2 2^{2-d}$ corresponding to two lines being internal, two being external, etc. The propagator for internal lines is the inverse of the coefficient of $\frac{1}{2}y^2$ in (6.6), namely $(2 + 2r)^{-1}$. It is easy to see that the absence of terms odd in y in (6.6) corresponds to rule 3 – removing vertices with an odd number of internal lines. Without rule 3 one would have had $\exp[-y^2 - Q(y + 2^{1-d/2}x)]$ instead of the exponential with two Q terms, in eq. (6.5). Then the ry^2 term in $Q(y)$ becomes $r(y + 2^{1-d/2}x)^2$ which has a cross term rxy. This is the term corresponding to the graph of fig. 6.2 which is excluded in the exponent of eq. (6.6) and which is not present in the complete theory. The variables x and y are analogous to the variables σ_{0q} and σ_{1q} in the theory of section 4; there is no term of the form $\sigma_{0q}\sigma_{1-q}$ in section 4.

The denominator in (6.5) removes all graphs without any external lines ("vacuum to vacuum" graphs in field theoretic terminology). Taking the logarithm of the right-hand side of (6.5) removes all disconnected graphs. So $-Q'(x)/2^d$ is the sum of all connected graphs with at least one external line. Note that the x^2 term in (6.6) passes through the integral (6.5) and contributes $4r$ to r'. We recognize this as the usual free field term which contributes $4r$ to r'.

The most important fact about (6.5) is that it produces the same set of graphs (apart from the restriction of rule 3) that are generated by expanding the exact functional integral. This is because the graphs themselves do not distinguish whether the integration variable is a single discrete variable (y) or a set of variables, or a function variable σ_q. The distinction is made only in terms of the rules for graphs: in the exact case a four-point vertex stands for $u_4(q \ldots q_3)$, for example, while in the approximate recursion formula the four point vertex stands for u. Furthermore, the topological factor associated with a graph (the factor 36 for the graph of fig. 6.1, for example) is also the same; it comes from the number of topologically equivalent ways of combining a given set of vertices and propagators and this is a counting problem which involves only the graphs themselves and not what they represent.

6.2. *Some numerical results*

Verification of the first two rules is now a matter of bookkeeping. One must show that the factors of P and p_0^2 produced by rules 1 and 2 can be absorbed into the coefficients c_1 and c_2 for every graph in the theory. Consider a general graph of l internal lines, $2m$ external lines and n vertices. In the exact theory, this graph has $n-1$ delta functions affecting the momenta of internal lines, and consequently $l - (n - 1)$ momentum integrals. There will be $2m + 2l$ lines at vertices. Now one can record the factors (apart from the topological factor) for this diagram. There will be

1. $(-c_1)^n$ – a factor $-c_1$ for each vertex.
2. c_2^{2m+2l} – a factor c_2 for each line at each vertex.
3. P^{l-n+1} – a factor P for each integral.
4. $(p_0^2 + 2c_1c_2^2r)^{-l}$ – a propagator for each internal line.

We want the contribution of this diagram to a coefficient in the expansion (6.2). Therefore, we must divide by an overall factor of $-c_1$ and a factor c_2^{2m} (one factor of c_2 for each external line). There is finally a factor coming from the change of scale which is $\zeta^{2m}2^{-(2m-1)d} = 2^{2m-(m-1)d}$. ζ^2 is equal to 2^{2+d} since no graphs are permitted to contribute q^2 terms to $u_2(q)$ (due to rule 1). Collecting all factors gives,

$$(-1)^{n-1}c_1^n c_2^{2m+2l}P^{l-n+1}\frac{1}{(p_0^2+2c_1c_2^2r)^l}\frac{1}{c_1c_2^{2m}}2^{2m-(m-1)d}. \tag{6.7}$$

In order that the dependence on P and p_0^2 disappear, choose

$$c_1 = P, \qquad c_2 = p_0/(2P)^{1/2}. \tag{6.8}$$

Now (6.7) becomes,

$$(-1)^{n-1}\ 2^{2m-(m-1)d}/(2+2r)^l. \tag{6.9a}$$

Consider now the same graph in the expansion of the approximate recursion formula. For each external line there is a factor $2^{1-d/2}$. For each propagator there is a factor $(2+2r)^{-1}$. Combining these factors for the $2m$ external lines and l internal lines gives

$$2^{2m-md}(2+2r)^{-l}. \tag{6.9b}$$

Besides these factors there is a factor $-u$ for each 4-point vertex, a factor $-w$ for each 6-point vertex, etc.; these factors also appear in the original interpretation of the graph except for the $-$ signs. There are also combinatorial factors such as how many ways the internal lines can be assigned at a vertex (the factor 6 in $6ux^2y^2 2^{2-d}$ in eq. (6.6) is a combinatorial factor) or the number of ways of hitching different vertices to form a given graph. The combinatorial factors are the same for both the functional integral expansion and the expansion of the recursion formula. There is one final factor: the graphs for the integrals on the right-hand side of (6.5) give $-Q'(x)/2^d$, not $Q'(x)$. So to obtain $Q'(x)$ one multiplies by -2^d. This changes the factor (6.9b) to

$$(-1)^{n-1}2^{2m-(m-1)d}(2+2r)^{-l},$$

which is identical to (6.9a).

This completes Polyakov's derivation [116] of the approximate recursion formula. Note that $\eta = 0$ for the approximate recursion formula; this follows from eq. (5.36) since $\zeta = 2^{d+2}$. The reason for this is that all momentum dependence of graphs has been ignored.

The approximate recursion formula is not a quantitative approximation to the exact renormalization group equations. So far as is known, it is not the first term in any exact expansion of the exact renormalization group equations. Nor is it the first term in a sequence of successive approximations converging to the exact equations. One can only judge its validity by comparing solutions of the approximate recursion formula with other calculations such as high temperature expansions (see sections 7 and 8). In this respect it does rather well.

Baker [117] and Dyson [118] have constructed models for which the recursion formula is exact. However, these models contain long-range interactions of a peculiar sort. Golner [119] has derived a modified form of the recursion formula in which η is not automatically zero.

The original derivation of the recursion formula [111] using "phase space cell analysis" is more compelling than Polyakov's derivation presented here. Phase space cell analysis is not discussed in these sections because it has already been described several times: thoroughly in [111], briefly in [120]. One should find these references less difficult to read after studying this section and

section 7. Phase space cell analysis was originally developed in connection with fixed source models of the nucleon [121]. The real importance of phase space analysis is that it is an idea one can use to get started on a problem for which all conventional techniques have failed. That in fact was the situation in critical phenomena at the time [111] was written. Another example of the use of phase space cell analysis is given by Langer and Bar-on [122].

We shall now discuss how one uses the recursion formula to study critical behavior without expanding in powers of u, w, etc. Since (6.5) is a non-linear integral equation, the only adequate tool to analyze it is the computer. Luckily, one dimensional integrals can be evaluated with great ease and accuracy numerically. The analysis of (6.5) begins with a search for the fixed point [111]. Consider an initial interaction with $Q_0(y) = r_0 y^2 + \frac{1}{2} y^4$. The coefficient r_0 is a parameter which one varies in order to locate the critical temperature. To do this one generates on the computer a sequence

$$Q_0(y) \to Q_1(y) \to \ldots \to Q_l(y) \to \tag{6.10}$$

for a given r_0. This sequence is examined for a fixed value of y, say $y = 1.5$. What one wants is a sequence tending to a limit for $l \to \infty$, as depicted in fig. 6.7. This never happens in practice. If r_0

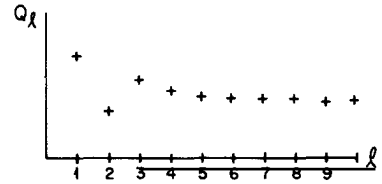

Fig. 6.7. An ideal plot of the approach of Q_l to a fixed point.

is chosen too large, the sequence goes off to infinity because one is above the critical point and the Gaussian term $r_l y^2$ dominates for large l, and then $r_{l+1} = 4r_l$. To compensate, one lowers r_0 and generates a new sequence Q_l. If r_0 were chosen too low, the sequence will tend to oscillate badly for large l. In practice one compromises between these two extremes by choosing an r_0 such that Q_l falls within a preassigned range after a preassigned number l' of iterations. This procedure is then repeated for larger values of l'. Once one has a sequence with $Q_l(y)$ stabilized over many iterations l' ($l' \geqslant 12$ in practice), one has a good approximation to the fixed point function Q^*.

In three dimensions the resulting function is shown in fig. 6.8. Further properties of the recursion formula and the fixed point function Q^* will be discussed in the next section.

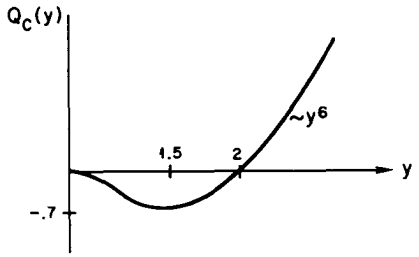

Fig. 6.8. Fixed point function $Q^*(y)$ for dimension 3 as determined numerically from the approximate recursion formula.

7.

This section is divided into two parts. First is a continuation of the presentation of the results of investigations using the approximate recursion formula. Then comes the derivation of a scaling formula for the n-spin correlation functions. This formula will allow us to compute ϵ expansions for critical exponents using Feynman graph techniques.

7.1. *More results from the approximate recursion formula*

Recall that last section's presentation of the approximate recursion formula began with the interaction

$$\mathcal{H}_0 = -\tfrac{1}{2} \int_x [\nabla s(x)]^2 - c_1 \int_x Q_0[c_2 s(x)] \tag{7.1}$$

where the possible momenta in this physical system run over the interval $0 < |q| < 1$. The recursion formula then generates a new interaction which describes the same physics for a system with a momentum cutoff one-half as large. In particular,

$$\mathcal{H}_l = -\tfrac{1}{2} \int_x [\nabla s(x)]^2 - c_1 \int_x Q_l[c_2 s(x)] \tag{7.2}$$

generates the same physics as (7.1) but with a momentum cutoff $|q| < 2^{-l}$, provided that

$$Q_{l+1}(y) = -2^d \ln\{I_l(2^{1-d/2}y)/I_l(0)\} \tag{7.3}$$

where

$$I_l(z) = \int_{-\infty}^{\infty} \exp[-y^2 - \tfrac{1}{2}Q_l(y+z) - \tfrac{1}{2}Q_l(-y+z)]\, dy. \tag{7.4}$$

These formulae have been studied numerically for several choices of Q_0 and dimensionality d of the physical system. Consider in particular

$$Q_0 = r_0 y^2 + u_0 y^4 \tag{7.5}$$

for d between 2 and 4. As discussed at the end of last lecture, given u_0 one adjusts r_0 to find the critical value, i.e. the value r_{0c} of r_0 for which,

$$Q_l(y) \to Q^*(y), \qquad \text{as } l \to \infty. \tag{7.6}$$

For $d = 3$ and $u_0 = 0.5$ one finds the fixed point function $Q^*(y)$ shown in fig. 7.1. For dimensions near 4, say 3.9 or 3.8 (and smaller u_0), one finds flatter curves as shown in the same figure. Recall from our earlier discussions that the fixed point function is trivial (i.e. $Q^* = 0$) in four dimensions, so the behavior of the curves in fig. 7.1 as d tends to four is expected. When d is precisely 4, and $u_0 < 0.35$, one finds that $u_l \sim 1/l$ for large l. So, free field behavior is approached very slowly as expected from other formulations of the renormalization group [123], (see section 13). In practice this means that when d is near four it takes many iterations of (7.3) and (7.4) in order to reliably map out the fixed point function.

As d varies from 3 to 2, Q^* becomes much more complicated. For $d = 2$, Q^* behaves roughly as shown in fig. 7.2. (This figure was obtained from a slightly modified form of (7.3). Namely,

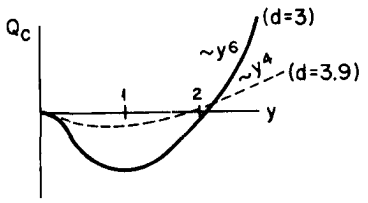

Fig. 7.1. Fixed point functions in dimension 3 and 3.9.

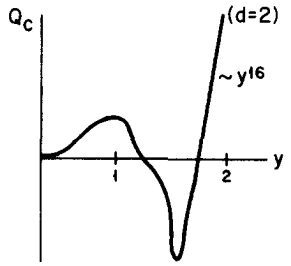

Fig. 7.2. Fixed point function in 2 dimensions. This curve resulted from a numerical study of a slightly modified form of the recursion formula.

the factor $2^{1-d/2}$ was replaced by $2^{1-d/2-1/8}$.) It is easily seen that two dimensions is rather special: Suppose one is calculating $Q_{l+1}(y)$ for large y. If $d = 2$, $Q_{l+1}(y)$ is determined by $I_l(y)$, i.e. the arguments of the two functions match. The behavior of $I_l(y)$ for large y is in turn determined by Q_l when its argument is large and comparable in size to y. Therefore, the asymptotic behavior of the input function Q_0 to the iteration formula will tend to propagate through many iterations. This behavior is contrary to the existence of a fixed point independent of Q_0, and makes it hard for a fixed point to occur. In contrast, for $d > 2$ the factor $2^{1-d/2}$ means that $Q_{l+1}(y)$ is determined by $Q_l(y')$ for $y' \sim 2^{1-d/2}y$ which is considerably less than y when y is large. Then there is no problem.

Once one has found the critical function $Q^*(y)$, further numerical studies of the recursion formula yield the critical exponent ν. For this purpose, one begins the iteration scheme with $r_0 \approx r_{0c}$, but not equal to r_{0c}. One then observes how $Q_l(y)$ departs from Q^*. For reasonably large l, one expects that Q_l has the form (by analogy with eq. (4.41))

$$Q_l \to Q^* + (r_0 - r_{0c})\lambda^l R^*(y) \tag{7.7}$$

where $R^*(y)$ is a function independent of l and r_0. For this to be true, l must be large enough so that initial transients have disappeared; then $r_0 - r_{0c}$ must be small enough so that $(r_0 - r_{0c})\lambda^l$ is small. λ is the leading eigenvalue of the linearized renormalization group equation as in section 4. The exponent ν is,

$$\nu = \ln 2/\ln \lambda \tag{7.8}$$

as in section 4. The eigenvalue λ can be determined by fitting (7.7) to the results of the numerical iterations.

In fig. 7.3 we plot the calculated values of ν as a function of d obtained from the approximate recursion formula and compare the results with the high temperature expansion result at $d = 3$ and the exact Onsager solution at $d = 2$. The calculated curve is fit well by a quadratic function of $\epsilon = 4 - d$.

Another use of the recursion formula is in studying the ϵ expansion. One can compute the ϵ expansion for ν resulting from the recursion formula by numerical techniques. The recursion formula yields the following series for 2ν:

$$2\nu = 1 + 0.167\epsilon + 0.04\epsilon^2 - 0.016\epsilon^3 + 0.077\epsilon^4 - 0.2\epsilon^5 + 0.67\epsilon^6 - 2.5\epsilon^7 + 10.3\epsilon^8 + \ldots \tag{7.9}$$

indicating that the series is really an asymptotic expansion. If one truncates the series at second order, one finds $\nu = 0.603$ at $d = 3$. This is to be compared with the best numerical result of the

recursion formula, $\nu = 0.609$. Keeping only the linear term in (7.9) gives $\nu = 0.583$. These results suggest that it is sensible to terminate the series at ϵ^2. A Padé analysis of (7.9) also reveals that there is no appreciable improvement over the ϵ^2 approximation until eighth order at which point the Padé expansion yields a slightly better value. This suggests that it is important to calculate the exact expansion for ν to order ϵ^2, but that there is not much point in calculating the ϵ^3 term or higher. The most important result from the recursion formula is the hint that the series to order ϵ^2 should give a good value for ν for $\epsilon = 1$. One might not otherwise have expected the series to be good for this large value of ϵ.

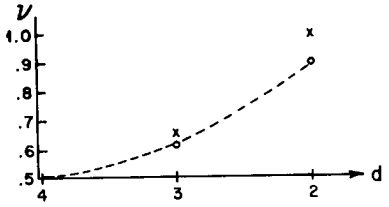

Fig. 7.3. Critical exponent ν plotted against the dimension of the physical system. The open circles resulted from numerical studies of the recursion formula. The x at $d = 2$ is the value of ν from the exact Onsager solution while the x at $d = 3$ is the best number available from high temperature expansions.

The next topic to be discussed is a method of calculating the ϵ expansion exactly to any order using Feynman graph techniques. It is first necessary, however, to establish a theorem concerning the correlation functions. The proof will require the rest of this lecture and the theorem will underlie much of the work to follow. Consider the correlation function defined for the original Hamiltonian $\mathcal{H}_0$, depending on temperature T,

$$\Gamma(q_1, \ldots, q_n; T) = Z^{-1} \langle \sigma_{q_1} \ldots \sigma_{q_n} \exp(\mathcal{H}_0) \rangle / \delta^{(d)}(q_1 + \ldots + q_n). \tag{7.10}$$

The field theory analogs of (7.10) are the vacuum expectation values of the products of n fields in momentum space. See section 10. We will show that if all the q_i are very small ($\ll 1$), and $T \approx T_c$, then

$$\Gamma(q_1, \ldots, q_n; T) = \xi^{(n-1)d - nd_s} \zeta_A^n F(\xi q_1, \ldots, \xi q_n), \tag{7.11}$$

where ξ is the correlation length and

$$d_s = \tfrac{1}{2}(d - 2 + \eta), \tag{7.12}$$

where η is a critical exponent. In field theory, d_s becomes the anomalous dimension [132] (in mass units) of the spin field. It will be calculated in a later lecture using Feynman graphs. The correlation length ξ depends upon $T - T_c$ and it is only through this factor and the normalization constant ζ_A that (7.11) contains any dependence on the original Hamiltonian.

Eq. (7.11) is true independently of the size of ξq_i. (ξ is large for $T \approx T_c$; one can make ξq_i small or large without violating the requirements $q_i \ll 1$, $T \approx T_c$.)

Eq. (7.11) is an example of the "scaling laws" proposed several years ago for critical phenomena

(see the reviews in [124]); this particular formula was proposed by Patashinskii [125]. None of the scaling laws have been proven from scratch for three dimensions and none will be proven here. Here it is assumed that there is a renormalization group transformation which has a fixed point and a simple form near the fixed point as exemplified by the approximate equations of sections 4 and 6. A proof from scratch would have to prove this assumption.

To obtain the result (7.11) requires the use of an exact renormalization group equation. For the present purpose it is convenient to use a differential form of the renormalization group transformation. In previous lectures the renormalization group transformation was a discrete transformation resulting from decreasing the cutoff from Λ to $\Lambda/2$. A differential transformation results from changing the cutoff infinitesimally, from Λ to $\Lambda - \delta\Lambda$. An explicit differential transformation will be constructed in section 11; for now it is sufficient to know that a differential transformation exists. In previous lectures the interaction with cutoff $\Lambda = 2^{-l}$ was denoted $\mathcal{H}_l$; here it is convenient to write $\Lambda = e^{-t}$ and denote the effective interaction with cutoff Λ by $\mathcal{H}_t$. In previous lectures the interaction $\mathcal{H}_{l+1}$ was determined by a recursion formula in terms of $\mathcal{H}_l$. The differential analogue to a recursion formula reads

$$d\mathcal{H}_t/dt = U[\mathcal{H}_t]. \tag{7.13}$$

An explicit form for the infinitesimal transformation U is given in section 11. The interaction $\mathcal{H}_t$ is a functional of the variable σ'_q, where

$$\sigma_q = \zeta(t, T)\sigma'_{q'}, \qquad q' = e^t q, \qquad 0 < |q'| < 1 \tag{7.14}$$

and $\zeta(t, T)$ is a renormalization parameter; this is a straightforward generalization from the discrete recursion scheme.

A crucial feature of eq. (7.13) is that $U[\mathcal{H}_t]$ has no explicit t dependence, i.e., it is not $U[\mathcal{H}_t, t]$. This is a generalization of the fact that the recursion formula for $\mathcal{H}_{l+1}$ does not depend on l.

As in section 3, one can compute $\Gamma(q_1, \ldots, q_n, T)$ for small enough q_i using $\mathcal{H}_t$ in place of $\mathcal{H}_0$. The result is,

$$\Gamma(q_1, \ldots, q_n, T) = Z_t^{-1}[\zeta(t, T)]^n \langle \sigma'_{e^t q_1} \ldots \sigma'_{e^t q_n} \exp\{\mathcal{H}_t[\sigma', T]\}\rangle / \delta^d(q_1 + \ldots + q_n) \tag{7.15}$$

provided

$$0 < |e^t q_i| < 1 \tag{7.16}$$

for all i. (The reader is warned that when $\mathcal{H}_t$ is defined using the explicit formulae of section 11, eq. (7.15) is replaced by a somewhat more complex formula for technical reasons. The theorem (eq. (7.11)) is unchanged and the argument given here is still correct in essence.) Z_t is the partition function $\langle \exp(\mathcal{H}_t)\rangle$.

To proceed one must know how $\zeta(t, T)$ changes with the cutoff. Recall that $\zeta(t, T)$ is determined by the requirement that the q^2 term in $u_2(q, t)$ in $\mathcal{H}_t$ have coefficient 1. Now, in the iterative form of the renormalization group, a single iteration determines the ratio ζ_{l+1}/ζ_l where ζ_l is the renormalization factor relating σ_q to $\sigma'_{2^l q}$ after l iterations. In a differential formulation one determines the ratio $\zeta(t + dt)/\zeta(t)$, i.e. $(1 + [\zeta(t)]^{-1}[d\zeta/dt]\, dt)$. This ratio is determined purely by $\mathcal{H}_t$: if the q^2 term in $u_2(q, t)$ has coefficient 1, $\zeta^{-1} d\zeta/dt$ must be chosen so that the q^2 term in $u_2(q, t + dt)$ has coefficient 1. So there will be an equation

$$\frac{1}{\zeta(t)}\frac{d\zeta}{dt} = V[\mathcal{H}_t] \tag{7.17}$$

where V is an unspecified functional of $\mathcal{H}_t$.

We want to consider the physical system only near the critical point. From experience with the recursion formula in previous lectures one already has a good idea of the behavior (7.13) will produce. In particular, suppose one could plot $\mathcal{H}_t$ (actually a term of $\mathcal{H}_t$ such as u_t) versus t: see fig. 7.4. For small t (relatively large cutoffs) $\mathcal{H}_t$ passes through a transient region until it reaches its critical value. The system is not quite at the critical point so eventually $\mathcal{H}_t$ departs from its critical value. $\mathcal{H}_t$ may go to zero or it may diverge as shown in fig. 7.4. The onset of this region is characterized by a value of t, t_0 say. The area to the right of t_0 we call the "correlation length region". Now imagine changing the parameters in $\mathcal{H}_0$ so that the differential equation (7.13) starts out closer to criticality. Then the instabilities in (7.13) do not appear until larger values of t.

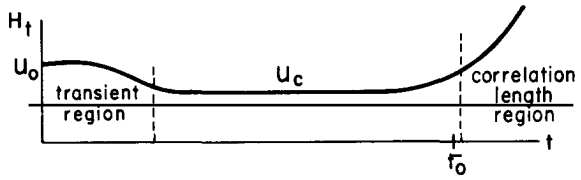

Fig. 7.4. $\mathcal{H}_t$ plotted against t. t is related to the cutoff Λ_t in $\mathcal{H}_t$ by $\Lambda_t = e^{-t}$.

However, after the transient region, memory of the initial $\mathcal{H}_0$ has been lost, so the right end of the new curve will be identical in shape to the old, but will simply be shifted to the right. To see this formally from the differential equation, suppose one has a solution $\mathcal{H}(t, T)$ depending on the temperature T. Consider the region of t somewhat less than t_0, where $\mathcal{H}(t, T)$ is departing from the fixed point $\mathcal{H}^*$, but not by much. In this range of t one expects $\mathcal{H}(t, T)$ to have the form

$$\mathcal{H}(t, T) = \mathcal{H}^* + (T - T_c)\, e^{at}\mathcal{H}_d \tag{7.18}$$

where a is the leading eigenvalue of the linearized renormalization group equation, and $\mathcal{H}_d$ is a fixed interaction. The eigenvalue λ of the discrete case (see, e.g., eq. (7.7)) corresponds to 2^a. The important feature of eq. (7.18) is that both the t and T dependence of $\mathcal{H}(t, T)$ are contained in the single parameter $(T - T_c)\, e^{at}$. In consequence, a change of T in the curve $\mathcal{H}(t, T)$ is equivalent to a translation in t, as claimed, as long as (7.18) is valid. One expects, as in previous discussions, that when e^t (the inverse of the cutoff) is of order ξ, then $\mathcal{H}(t, T)$ will be appreciably different from $\mathcal{H}^*$, i.e. $(T - T_c)\, e^{at} \sim 1$ when $e^t \sim \xi$. This means that a is $1/\nu$. This also means that for $e^t > \xi$, $\mathcal{H}(t, T)$ is no longer near $\mathcal{H}^*$ and the linear approximation (7.18) is invalid. Further analysis allows us to treat this case. Define $t_0(T)$,

$$(T - T_c) = \exp\{-t_0(T)/\nu\}. \tag{7.19}$$

Then,

$$\mathcal{H}(t, T) = \mathcal{H}^* + \exp\{[t - t_0(T)]/\nu\}\mathcal{H}_d \tag{7.20}$$

in the linear region. Thus it looks as if $\mathcal{H}(t, T)$ can be written

$$\mathcal{H}(t, T) = \mathcal{H}_s(t - t_0(T)) \tag{7.21}$$

where $\mathcal{H}_s(t)$ is a particular solution of the renormalization group equations. It does not depend on T and is defined for $-\infty < t < \infty$. Because (7.13) is independent of t, $\mathcal{H}_s(t - t_0)$ is a solution if $\mathcal{H}_s(t)$ is. This ensures that $\mathcal{H}_s(t - t_0(T))$ continues to be $\mathcal{H}(t, T)$ even if t is too large for the linearized equation (7.18) to hold. Hence, the "trajectories" $\mathcal{H}_t[\sigma', T]$ have identical shapes for

different values of T, as claimed. Only for small t, for which initial transients are present, is $\mathcal{H}(t, T)$ not equal to $\mathcal{H}_s(t - t_0(T))$.

A further word about the initial transient region (small t). There are many possible choices for the input interaction $\mathcal{H}_0$, in addition to the choice of the temperature T. There are many parameters in $\mathcal{H}_0$; if one is completely general the entire functions $u_2(q)$, $u_4(q \ldots q_3)$, etc. are variable. Only one parameter, typically $r_0 = u_2(0)$, has to be varied to find the critical point. For small t, $\mathcal{H}_t$ obviously will depend on all the parameters in $\mathcal{H}_0$. But if $r_0 = r_{0c}$, its critical value, then $\mathcal{H}_t \to \mathcal{H}^*$ for $t \to \infty$; and $\mathcal{H}^*$ is independent of the initial parameters in $\mathcal{H}_0$. This is because $\mathcal{H}^*$ is a solution of the renormalization group equations, and, at least for the examples discussed in sections 4 and 6, $\mathcal{H}^*$ is completely determined by the renormalization group equations: there are no free parameters in $\mathcal{H}^*$ so $\mathcal{H}^*$ cannot depend on any of the parameters in $\mathcal{H}_0$.

Likewise in the example discussed in section 4 there is only one growing solution of the linearized renormalization group equation, up to a multiplicative constant. But a multiplicative constant can be absorbed into a translation in t, so all the growing solutions can be obtained from a single solution $\mathcal{H}_s(t)$. $\mathcal{H}_s(t)$ is defined so that $\mathcal{H}_s(-\infty) = \mathcal{H}^*$; $\mathcal{H}_s(t)$ starts out at $t = -\infty$ at the fixed point. For large negative t, $\mathcal{H}_s$ has the linearized form,

$$\mathcal{H}_s(t) = \mathcal{H}^* + e^{t/\nu}\,\mathcal{H}_d\,. \tag{7.22}$$

For $t > 0$, $\mathcal{H}_s(t)$ can differ appreciably from $\mathcal{H}^*$ but is still unique.

The independence of $\mathcal{H}^*$ and $\mathcal{H}_s(t)$ of any parameters in $\mathcal{H}_0$ is the basis for universality. This will be discussed further in section 12. The solution $\mathcal{H}_s(t)$ is called a "renormalized trajectory": see section 12.

7.2. *Scaling theorems for n-spin correlation functions*

Now let us return to the analysis of the n-spin correlation function (7.15).

For large t one can replace $\mathcal{H}_t[\sigma', T]$ by $\mathcal{H}_s[\sigma', t - t_0(T)]$. Define,

$$f(q_1', \ldots, q_n', \tau) = \frac{\langle \sigma'_{q_1'} \ldots \sigma'_{q_n'} \exp\{\mathcal{H}_s[\sigma', \tau]\}\rangle}{\delta^d(q_1' + \ldots + q_n')\, Z(\tau)}\,. \tag{7.23}$$

Then,

$$\Gamma(q_1, \ldots, q_n; T) = \frac{\zeta^n(t, T) f(e^t q_1, \ldots, e^t q_n; t - t_0(T))\, \delta^{(d)}(e^t q_1 + \ldots + e^t q_n)}{\delta^d(q_1 + \ldots + q_n)} \tag{7.24}$$

where we have explicitly displayed the delta function in the numerator of (7.15). The delta functions cancel after producing an overall factor of e^{-dt},

$$\Gamma(q_1, \ldots, q_n; T) = \zeta(t, T)^n\, e^{-dt} f(e^t q_1, \ldots, e^t q_n; t - t_0(T)). \tag{7.25}$$

This relation holds only when:

1. t is outside the transient region, i.e. t large;
2. $e^t q_i < 1$ for all $i \leqslant n$.

This means that one must consider only very small values of q_i.

It is convenient to introduce the variable $\tau = t - t_0(T)$ and hold it fixed independent of T. Rewrite (7.20),

$$\Gamma(q_1, \ldots, q_n; T) = [\zeta(\tau + t_0(T), T)]^n\, e^{-d(t_0 + \tau)} f(e^{t_0 + \tau} q_1, \ldots, e^{t_0 + \tau} q_n; \tau). \tag{7.26}$$

This formula accomplishes the following. The T dependence was originally in the last argument of F (as in eq. (7.24)) where it is hard to control. Now the T dependence has been transferred to the explicit $\exp(t_0(T))$ factors which are more manageable. The factors $\exp(t_0(T))$ can be expressed in terms of the correlation length ξ. Consider for example the definition of ξ adopted in section 3:

$$\xi^2 = -\mathrm{d} \ln \Gamma(q)/\mathrm{d}q^2|_{q=0}. \tag{7.27}$$

Using eq. (7.26), one obtains

$$\xi^2(T) = \exp\{2(t_0(T) + \tau)\} \mathrm{d} \ln f(q', \tau)/\mathrm{d}q'^2|_{q'=0}. \tag{7.28}$$

A consequence of this formula is that one can write

$$\exp\{t_0(T)\} = \xi(T)/\bar{\xi}(\tau) \tag{7.29}$$

where $\bar{\xi}(\tau)$ is a function of τ and independent of T. It is a consequence of this formula that $\bar{\xi}(\tau)$ is independent of τ also.

Finally, one must determine the dependence of $\zeta(t, T)$ on t and T. This is straightforward given the differential equation (7.17). Over much of the range of t, $\mathcal{H}_t$ is approximately $\mathcal{H}^*$, so consider first the simpler equation $\mathrm{d} \ln \zeta/\mathrm{d}t = V[\mathcal{H}^*]$. It is convenient to define d_s as

$$d - d_s = V[\mathcal{H}^*] \tag{7.30}$$

(d_s is a constant independent of T and t).

A solution of the simple equation is

$$\zeta(t, T) = \exp\{(d - d_s)t\}. \tag{7.31}$$

There are two ranges of t for which the simplified equation is incorrect; the formula (7.31) will be modified to take this into account. First of all, for $t \gtrsim t_0$, $\mathcal{H}_t$ is not close to $\mathcal{H}^*$; but for this range of t, $\mathcal{H}_t$ can be replaced by $\mathcal{H}_s[\sigma', t - t_0(T)]$. This means $V[\mathcal{H}_t]$ is a function only of $t - t_0(T)$, say $v(t - t_0(T))$. For $t \ll t_0(T)$, $v(t - t_0(T))$ reduces to $d - d_s$. Taking the function v into account, one can write

$$\mathrm{d} \ln \zeta/\mathrm{d}t = d - d_s + [v(t - t_0(T)) - (d - d_s)]. \tag{7.32}$$

This is to be integrated with the boundary condition $\zeta = 1$ for $t = 0$. The quantity $t_0(T)$ is large for the range of T near T_c of interest, and for $t \ll t_0(T)$ the term in brackets is negligible. Thus when (7.32) is integrated one is free to integrate the bracketed term from $t = -\infty$ to t instead of 0 to t. Then one has,

$$\ln \zeta(t, T) = (d - d_s)t + \ln \zeta_B(t - t_0(T)) \tag{7.33}$$

where

$$\ln \zeta_B(\tau) = \int_{-\infty}^{\tau} [v(\tau_1) - (d - d_s)] \mathrm{d}\tau_1. \tag{7.34}$$

One now has,

$$\zeta(t, T) = \zeta_B(t - t_0(T)) \exp\{(d - d_s)t\}. \tag{7.35}$$

(Note that ζ_B is 1 when $t \ll t_0(T)$.)

Finally, there is a correction to this formula due to the transient region for small t. As long as

$T \approx T_c$, the behavior of $\mathcal{H}_t$ for small t is independent of both T and the final large value of t. In consequence one finds that the correct form for $\ln \zeta$ is,

$$\ln \zeta(t, T) = (d - d_s)t + \ln \zeta_B[t - t_0(T)] + \ln \zeta_A \tag{7.36}$$

where ζ_A is independent of both t and T (provided t is large and $T \approx T_c$). The constant ζ_A does depend on variables (besides T) that appear in $\mathcal{H}_0$, such as u_0 in section 4. Using eqs. (7.26) and (7.36) one has

$$\Gamma(q_1, \ldots, q_n, T) = [\zeta_A \zeta_B(\tau) \exp\{(d - d_s)\tau\}]^n \, e^{-d\tau} \left[\frac{\xi(T)}{\bar{\xi}}\right]^{(n-1)d - nd_s} f(e^\tau \xi q_1/\bar{\xi}, \ldots, e^\tau \xi q_n/\bar{\xi}, \tau). \tag{7.37}$$

To obtain eq. (7.11) one defines

$$F(q'_1, \ldots, q'_n, \tau) = [\zeta_B(\tau) \exp(-d_s\tau) \, \bar{\xi}^{d_s}]^n \left[\frac{e^{d\tau}}{\bar{\xi}^d}\right]^{n-1} f(e^\tau q'_1/\bar{\xi}, \ldots, e^\tau q'_n/\bar{\xi}, \tau). \tag{7.38}$$

Then,

$$\Gamma(q_1, \ldots, q_n, T) = [\xi(T)]^{(n-1)d - nd_s} F(\xi q_1, \ldots, \xi q_n, \tau) \, \zeta_A^n. \tag{7.39}$$

But since neither Γ nor ξ depend on τ, F must also be independent of τ; thus one obtains eq. (7.11).

A stronger result can be obtained for momenta large compared to $1/\xi$ but small compared to 1. Return to eq. (7.15) and consider the case that t lies in the range $t \gg 1$, $t_0(T) - t \gg 1$ so that $\mathcal{H}_t \approx \mathcal{H}^*$. Define,

$$f^*(q'_1, \ldots, q'_n) = (Z^*)^{-1} \langle \sigma'_{q'_1} \ldots \sigma'_{q'_n} \exp\{\mathcal{H}^*[\sigma']\}\rangle / \delta^d(q'_1 + \ldots + q'_n). \tag{7.40}$$

For t in this range ζ_B is 1; therefore

$$\Gamma(q_1, \ldots, q_n, T) = \zeta_A^n \exp\{n(d - d_s)t\} \, e^{-dt} f^*(e^t q_1, \ldots, e^t q_n). \tag{7.41}$$

This equation has two consequences. First, it has no T dependence on the right-hand side, which means one can replace T by T_c:

$$\Gamma(q_1, \ldots, q_n, T) = \Gamma(q_1, \ldots, q_n, T_c). \tag{7.42}$$

The only apparent restriction on this equation is (7.16); since t must be large, the restriction is that $|q_i| \ll 1$. However, the $\mathcal{H}_d$ term in eq. (7.18) has been neglected in replacing $\mathcal{H}_t$ by $\mathcal{H}^*$ and this is legitimate only if $\xi^{-1} \ll |q_i|$ for all $|q_i|$ (see below). The second consequence is a scaling law which will be stated for $T = T_c$:

$$\Gamma(q_1, \ldots, q_n, T_c) = s^{(n-1)d - nd_s} \Gamma(sq_1, \ldots, sq_n, T_c) \tag{7.43}$$

for an arbitrary scale factor s, provided both $|q_i|$ and $s|q_i|$ are $\ll 1$. (To prove eq. (7.43), calculate $\Gamma(sq_1, \ldots, sq_n, T_c)$ using (7.41) with q_i replaced by sq_i and t replaced by $t' = t - \ln s$.) The scaling law (7.43) is (to field theorists) a consequence of scale invariance for vacuum expectation values of a field with anomalous dimension d_s. To see this more clearly, let $\Gamma(x_1, \ldots, x_n, T_c)$ be the coordinate space multi-spin correlation functions

$$\Gamma(x_1, \ldots, x_n, T_c) = \int_{q_1} \ldots \int_{q_n} \exp\{iq_1 \cdot x_1 + \ldots + iq_n \cdot x_n\} \, \delta^d(q_1 + \ldots + q_n) \Gamma(q_1, \ldots, q_n, T_c). \tag{7.44}$$

The scaling law in coordinate space analogous to (7.43) is

$$\Gamma(sx_1, \ldots, sx_n, T_c) = s^{-nd_s}\Gamma(x_1, \ldots, x_n, T_c) \tag{7.45}$$

and is valid for $|x_i| \gg 1$. In an exactly scale invariant field theory there would be no such restriction; the restriction occurs because the statistical mechanics has a cutoff momentum ($|q_i| \leqslant 1$) and therefore a cutoff length ~ 1; scale invariance only occurs at distances large compared to the cutoff length. The relation of d_s to η (eq. (7.12)) follows from comparison of eq. (7.45) with $n = 2$ with eq. (2.4) of section 2.

Now the restriction $\xi^{-1} \ll |q_i|$ for eq. (7.42) to hold will be explained. In order to neglect $\mathcal{H}_d$ in eq. (7.18) the size of $(T - T_c)e^{at}\mathcal{H}_d$ must be small. One can choose t and $T - T_c$ so that $(T - T_c)e^{at}$ is small; the problem is the "size" of $\mathcal{H}_d$. $\mathcal{H}_d$, like any interaction, must be the integral of an energy density:

$$\mathcal{H}_d = \int_x \mathcal{H}_d(x) \tag{7.46}$$

and the "size" of $\mathcal{H}_d$ depends on how large a volume is important for the integral. The approximate recursion formula provides an example of the form (7.46): within this approximation, $\mathcal{H}_d$ is

$$\mathcal{H}_d \propto -c_1 \int_x R^*[c_2 s(x)] \tag{7.47}$$

(cf. (7.2) and (7.7)). Consider the x-space form of the two-spin correlation function for $\mathcal{H}_{t_1}$,

$$\Gamma(x_1) = \langle s(x_1)s(0) \exp(\mathcal{H}_{t_1})\rangle. \tag{7.48}$$

This function behaves for very large x_1 like $\exp(-|x_1|/\xi_1)$ where ξ_1 is the correlation length for $\mathcal{H}_{t_1}$. Clearly ξ_1 can be ignored only when $|x_1| \ll \xi_1$, and ξ_1 becomes infinite if $\mathcal{H}_{t_1}$ is replaced by $\mathcal{H}^*$. The region over which $\mathcal{H}_d(x)$ is important for $\Gamma(x_1)$ is expected to be a region of size several times $|x_1|$ surrounding the points x_1 and 0, from rough locality considerations. If $|x_1| \ll \xi_1$, ξ_1 is negligible in $\Gamma(x_1)$ which suggests this region is small enough so the $\mathcal{H}_d$ term in (7.18) is only a small perturbation. For $|x_1| \gtrsim \xi_1$ the region is large enough so $\mathcal{H}_d$ is no longer a perturbation.

The general rule in coordinate space is that the $\mathcal{H}_d$ term is negligible only for correlation functions at distances short compared to the correlation length. This condition translated into momentum space means all the $|q_i|$ must be much larger than ξ^{-1}. This was a very cursory argument; a more careful but not rigorous analysis gives the same result.

This discussion of when $\mathcal{H}_d$ can be neglected can be avoided by assuming that $\Gamma(q_1, \ldots, q_n, T)$ has a well-defined limit for $T \to T_c$, e.g. $\Gamma(q_1, \ldots, q_n, T)$ becomes independent of $\xi(T)$ for $T \to T_c$. Then the scaling law (7.43) follows directly from (7.11). The reason for discussing the strength of $\mathcal{H}_d$ is that this is an example of a very general problem in renormalization group theory. The general problem is this: One has an interaction $\mathcal{H}_A + \theta\mathcal{H}_B$ where θ is a small parameter and $\mathcal{H}_A$ and $\mathcal{H}_B$ are two arbitrary interactions. Under what circumstances can $\mathcal{H}_B$ be neglected? For example, $\theta\mathcal{H}_B$ might be a transient term for large t, which has a small coefficient; we have consistently neglected such terms. According to the above discussion it should be legitimate to neglect $\mathcal{H}_B$ if one is calculating spin-spin correlation functions with $|x_i| \sim 1$. However, for correlation functions at large x, one has to consider the possibility that $\mathcal{H}_B$ is enhanced due to large volume effects and is no longer a perturbation.

A warning is in order here. Suppose one takes a short distance correlation function and differentiates it with respect to a parameter in the interaction, such as the temperature. For illustration let the parameter be θ. The resulting expression can be very sensitive to the $\mathcal{H}_B$ term even if θ is small. The reason is that after differentiation one has a new correlation function containing $\mathcal{H}_B = \int_x \mathcal{H}_B(x)$ integrated over all space: the new correlation function is no longer a short distance correlation function.

7.3. *Slow transients and their removal for small ϵ*

Now another problem also needed as background for the ϵ expansion calculation will be considered.

A problem that arises when d is near 4 is the slow approach of the effective interactions $\mathcal{H}_t$ (or $\mathcal{H}_l$ in the discrete case) to $\mathcal{H}^*$ when one is at the critical temperature. This slow approach can be illustrated using the simplified recursion formula of section 4. The study of this problem is important for the calculations in the next section.

The simplified recursion formulae (4.27) have a fixed point (r^*, u^*) given by (4.30). Suppose the initial value of u_0 is close to u^* but not equal to it, and that $r_0 = r_{0c}(u_0)$. Then the solution (r_l, u_l) of the recursion formula approaches (r^*, u^*) for $l \to \infty$. The rate at which r_l and u_l approach r^*, u^* can be determined by solving the linearized equations for $r_l - r^*$ and $u_l - u^*$. This was done in section 4. The result is that $r_l - r^*$ and $u_l - u^*$ behave as $2^{-\epsilon l}$ for large l, where $2^{-\epsilon}$ is the smaller eigenvalue of the linearized equation. When ϵ is small $u_l - u^*$ goes very slowly to 0; for $u_l - u^*$ to be very small, ϵl must be very large. In terms of the effective cutoff $\Lambda_l = 2^{-l}$, one must have $\epsilon \ln \Lambda_l$ large.

This slow approach is undesirable for the purposes of the following section. Fortunately, it can be avoided. If one chooses $u_0 = u^*$ exactly, then $u_l = u^*$ for all l and there is no problem with the slow approach anymore.

This is not the whole story. The full renormalization group recursion formula involves many other variables besides r_l and u_l. To simplify the discussion consider a single irrelevant variable w_l. It is not practical to set $w_0 = w^*$ because in practice there are too many irrelevant variables. Fortunately it is not necessary to put $w_0 = w^*$. Consider initial conditions (r_0, u_0, w_0) near the fixed point (r^*, u^*, w^*). Then the solution (r_l, u_l, w_l) can be computed by setting up and solving the linearized equations for $r_l - r^*$, $u_l - u^*$, and $w_l - w^*$. There will be three linearly independent solutions of the linearized equations corresponding to three different eigenvalues. In the limit $\epsilon \to 0$ these eigenvalues are 4, 1, and $\frac{1}{4}$ ($\frac{1}{4}$ is the largest eigenvalue for any irrelevant variable in the $\epsilon \to 0$ limit). For ϵ small, the third eigenvalue is still close to $\frac{1}{4}$.

The growing solution (eigenvalue near 4) is eliminated by choosing r_0 to be its critical value $r_{0c}(u_0, w_0)$. Similarly the solution with eigenvalue near 1 can be eliminated by proper choice of u_0: there will be a choice $u_0 = u_{0c}(w_0)$ such that $(r_l - r^*, u_l - u^*, w_l - w^*)$ involves only the third solution with eigenvalue near $\frac{1}{4}$. When w is neglected altogether, the "critical value" u_{0c} is just u^*. With w taken into account, $u_{0c}(w_0)$ is equal to u^* only if $w_0 = w^*$. But for other values of w_0, there is still a choice of u_0 which makes the coefficient of the slowly decreasing solution vanish. In fact, $u_{0c}(w_0)$ turns out to be u^* to order ϵ; only in order ϵ^2 is $u_{0c}(w_0)$ different from u^*. This might have been expected since the presence of w only changes u^* itself to order ϵ^2.

In conclusion, the initial values of all irrelevant variables can be chosen arbitrarily (typically they are set equal to zero). There will then be critical values r_{0c} and u_{0c} such that $\mathcal{H}_l \to \mathcal{H}^*$ like 4^{-l} instead of $2^{-\epsilon l}$. By setting $r_0 = r_{0c}$ and $u_0 = u_{0c}$, $\mathcal{H}_l$ becomes approximately equal to $\mathcal{H}^*$ after a few iterations regardless of how small ϵ is (r_{0c} and u_{0c}, like r^* and u^*, depend on ϵ).

8. Feynman graph calculation of critical exponents (ϵ expansion)

In this section we shall discuss the calculation of critical exponents in powers of ϵ using Feynman graph techniques [126]. This approach will be illustrated within the context of statistical mechanics although it will also serve as an introduction to following sections on field theory and anomalous dimensions. In particular, we wish to calculate the behavior of the spin-spin correlation function near the critical temperature. From the discussion of the last section (see eq. (7.45)) one expects that

$$\Gamma(x) \sim 1/x^{2d_s} \qquad \text{(denoted } \Gamma(x, 0) \text{ in section 7)} \tag{8.1}$$

for $1 \ll x \ll \xi$ and $T \approx T_c$. The calculation of this section will yield d_s, the anomalous dimension of the spin field, when the dimensionality d of the physical system is near four. Various critical exponents have also been calculated. See [126, 127], and table 8.1.

Table 8.1
Formulae obtained to date for critical exponents from Feynman graph calculations. See section 2 for definitions of exponents ($M = H^{1/\delta}$ at $T = T_c$ defines δ). The number n is the number of components of the spin s; elsewhere in these lectures $n = 1$. The constant T has been calculated to be 0.60103, approximately: see [127]. The formulae for η and γ are obtained in [127]. The other exponents are derived from scaling laws; the scaling laws are derived for the ϵ expansion in [131]:

$$\alpha = -\frac{(n-4)}{2(n+8)}\epsilon - \frac{(n+2)^2}{4(n+8)^3}(n+28)\epsilon^2 - \frac{(n+2)}{8(n+8)^5}\{n^4 + 50n^3 + 920n^2 + 3472n + 4800 - 192(5n+22)(n+8)T\}\epsilon^3$$

$$\beta = \tfrac{1}{2} - \frac{3}{2(n+8)}\epsilon + \frac{(n+2)(2n+1)}{2(n+8)^3}\epsilon^2 + \frac{(n+2)}{8(n+8)^5}\{3n^3 + 128n^2 + 488n + 848 - 48(5n+22)(n+8)T\}\epsilon^3$$

$$\gamma = 1 + \frac{(n+2)}{2(n+8)}\epsilon + \frac{(n+2)}{4(n+8)^3}(n^2 + 22n + 52)\epsilon^2 + \frac{(n+2)}{8(n+8)^5}\{n^4 + 44n^3 + 664n^2 + 2496n + 3104 - 96(5n+22)(n+8)T\}\epsilon^3$$

$$\delta = 3 + \epsilon + \frac{1}{2(n+8)^2}\{n^2 + 14n + 60\}\epsilon^2 + \frac{1}{4(n+8)^4}\{n^4 + 30n^3 + 276n^2 + 1376n + 3168\}\epsilon^3 + \frac{1}{16(n+8)^6}\{2n^6 + 96n^5$$
$$+ 1778n^4 + 12760n^3 + 50280n^2 + 147136n + 263040 + 768(n+2)(n+8)(5n+22)T\}\epsilon^4$$

$$\eta = \frac{(n+2)}{2(n+8)^2}\epsilon^2 + \frac{(n+2)}{8(n+8)^4}(-n^2 + 56n + 272)\epsilon^3 + \frac{(n+2)}{32(n+8)^6}\{-5n^4 - 230n^3 + 1124n^2 + 17920n + 46144$$
$$- 768(5n+22)(n+8)T\}\epsilon^4$$

$$\nu = \tfrac{1}{2} + \frac{(n+2)}{4(n+8)}\epsilon + \frac{(n+2)}{8(n+8)^3}(n^2 + 23n + 60)\epsilon^2 + \frac{(n+2)}{32(n+8)^5}\{2n^4 + 89n^3 + 1412n^2 + 5904n + 8640$$
$$- 192(5n+22)(n+8)T\}\epsilon^3$$

This section is rather brief, but there are now many other references on the ϵ expansion; see the supplemental list of references at the end of this report.

The basic idea of the Feynman graph calculation is the following: For physical systems with d near four and with scalar coupling $s^4(x)$, the strength u of the $s^4(x)$ interaction is of order $\epsilon = 4-d$ at the fixed point. Therefore, it is sensible to calculate perturbatively in the parameter u. If $d = 4$, one would be dealing with the usual Feynman graphs of ϕ^4 scalar field theory. Then, the Fourier transform of $\Gamma(x)$ becomes the propagator of the field theory and has the form (at T_c),

$$\Gamma(q) \propto \frac{1}{q^2}\{1 + a_1 u_0 \log q^2 + a_2 u_0^2 \log^2 q^2 + \ldots\}. \tag{8.2}$$

Here u_0 is the bare s^4 coupling constant and a_i are certain constants. The field theory is cutoff dependent; the cutoff has been set equal to unity as in previous lectures. For $\epsilon \neq 0$ all that happens is that more logarithms appear in order ϵ, ϵ^2, etc. See later in this section. To obtain useful information from the first few terms of (8.2), we must restrict our discussion to values of u_0 and q^2 such that

$$|u_0 \log q^2| \ll 1. \tag{8.3}$$

In principle one could relax (8.3) and sum the logarithms in (8.2) to all orders by other techniques. However, using what we have learned about the renormalization group and critical behavior, it will be sufficient to study only the first few terms in the expansion.

At the critical temperature one has from the Fourier transformation of (8.1),

$$\Gamma(q) \sim 1/q^{2-\eta} \tag{8.4}$$

($\eta = 2(d_s + 1) - d$). This equation can be expanded, if η is small:

$$\Gamma(q) \sim \frac{1}{q^2}\{1 + \eta \log q + \tfrac{1}{2}\eta^2 \log^2 q + \ldots\}. \tag{8.5}$$

If it were sensible to compare coefficients of the logarithms in (8.2) and (8.5), we could actually compute η directly. In general, matching power series (8.5) and (8.2) is not sensible. This is so because for most choices of u_0, the power law (8.4) is correct only when q^2 is so small that $u_0 \ln q^2$ is large, in which case the expansion (8.2) is invalid. This tight restriction on q^2 is due to the slow approach of $\mathcal{H}_l$ to $\mathcal{H}^*$ discussed in the previous section.

Fortunately it is possible to choose the parameters of the initial interaction such that eq. (8.4) is valid for a larger range of q^2. In terms of last section's discussion, one must choose $u_0 = u_{0c}(\epsilon)$ so that $\mathcal{H}_l$ settles down to its critical form $\mathcal{H}^*$ after only a few iteration steps. This means that (8.4) and (8.5) are valid for all q with $q \ll 1$, independently of ϵ. Then for small ϵ, which means small u_0, (8.2) and (8.5) are both valid for q in the range,

$$\exp(-1/u_0) \ll q \ll 1$$

and the two equations can be matched order by order in $\ln q^2$.

Now we turn to the calculation for $u_{0c}(\epsilon)$ (denoted $u_0(\epsilon)$ from now on) and η. To do this we shall study the following two quantities: the four point function u_R with its external legs set to zero momentum for $T \neq T_c$, and the propagator Γ with its external legs carrying momentum q for $T = T_c$. Then, if we define the propagator for $T \neq T_c$, but $q = 0$ (i.e., the susceptibility) to be r^{-1}, it follows from the scaling laws for n-point functions discussed last time that

$$u_R \propto r^{(\epsilon - 2\eta)/(2-\eta)} \tag{8.6}$$

as we shall show. A second expression for u_R follows from a Feynman graph expansion for u_R in powers of u_0. Matching the graph expansion with (8.6) determines $u_0(\epsilon)$. Then η is finally obtained in terms of ϵ by matching expansions (8.5) and (8.2).

Our first task is to prove (8.6) from the theorems developed in the last section. The initial interaction for the system will be,

$$\mathcal{H}_0 = -\tfrac{1}{2}\int_q [q^2(1+q^2)^2 + r_0]\,\sigma_q \sigma_{-q} - u_0 \int_{q_1}\int_{q_2}\int_{q_3} \sigma_{q_1}\sigma_{q_2}\sigma_{q_3}\sigma_{-q_1-q_2-q_3} \tag{8.7}$$

where q is allowed to range from zero to infinity. The kinetic energy piece of $\mathcal{H}_0$ has additional

q^2 dependence which serves as a cutoff in place of an upper limit on q. (Nitpickers will discover that one power of $1+q^2$ would have been sufficient.) This form of cutoff is more convenient for calculating Feynman graphs. It is also convenient to perform a "mass renormalization" on (8.7). Write

$$\mathcal{H}_0 = -\tfrac{1}{2} \int_q [q^2(1+q^2)^2 + r]\, \sigma_q \sigma_{-q} - \tfrac{1}{2} \int_q (r_0 - r)\, \sigma_q \sigma_{-q} - (u_0 \text{ term}). \tag{8.8}$$

The new term in (8.8), a self-mass correction, is treated as a perturbation as is the u_0 term. Since the r dependence is added and subtracted in (8.8), it can be chosen arbitrarily. We will choose it such that $\Gamma(0)$ ($q = 0$) is precisely r^{-1}, i.e. we demand that $\Gamma(0)$ be given as if only the first term in (8.8) were present; all corrections to $\Gamma(0)$ due to the perturbations must vanish identically.

Define the n spin correlation function in the usual way,

$$\Gamma(q_1, \ldots, q_n; r) = \frac{\langle \sigma_{q_1} \ldots \sigma_{q_n} \exp(\mathcal{H}_0) \rangle}{\delta^{(d)}(q_1 + \ldots + q_n)} \tag{8.9}$$

where we have indicated explicitly that Γ depends on r. The definition of u_R will be precisely,

$$u_R = \Gamma(0, 0, 0, 0; r)_{\text{connected}} / \Gamma(0, r)^4 \tag{8.10}$$

where the subscript "connected" means that no graphs consisting of the products of 2 point functions should appear in (8.10). The denominator $\Gamma^4(0, r)$ is equivalent to removing self-energy insertions on the external legs of all diagrams. To discover how u_R depends upon r we refer to the scaling laws obtained in the last lecture. In particular from (7.11) we have

$$\Gamma(q_1, \ldots, q_n; r) \sim \xi(r)^{(n-1)d - nd_s} F(q_1 \xi, \ldots, q_n \xi) \tag{8.11}$$

so,

$$\Gamma(n \text{ zeros}; r) \sim \xi(r)^{(n-1)d - nd_s}. \tag{8.12}$$

The propagator then has the property,

$$\Gamma(0; r) \sim \xi^{d - 2d_s} = \xi^{2-\eta}. \tag{8.13}$$

But, $\Gamma(0; r) = r^{-1}$ by definition, so r and ξ are simply related,

$$r^{-1} \sim \xi^{2-\eta}. \tag{8.14}$$

From (8.12) the correlation function for four spins satisfies,

$$\Gamma(0, 0, 0, 0; r) \sim \xi^{3d - 4d_s} = \xi^{d+4-2\eta}. \tag{8.15}$$

Therefore, combining (8.13) and (8.15),

$$u_R \sim \xi^{d+4-2\eta} / \xi^{8-4\eta}. \tag{8.16}$$

In terms of r this reads

$$u_R \sim r^{(\epsilon - 2\eta)/(2-\eta)} \tag{8.17}$$

as claimed.

Our next task is to obtain power series for u_R and $\Gamma(q; r)$ in the parameter u_0. This is an exercise in computing Feynman diagrams in 3.99 dimensions [126, 128]. The graphs contributing

to u_R to order u_0^3 are shown in fig. 8.1 and the graphs contributing to $\Gamma(q; r)$ to order u_0^2 are given in fig. 8.2. Before illustrating the calculation of a graph contributing to u_R, consider the graphs in fig. 8.2. Since the self-mass term in the Hamiltonian is defined in order that $\Gamma(0; r)$ be just the free propagator at $q = 0$ and since the loop integral in the second diagram in fig. 8.2 generates no q dependence, the second graph is completely cancelled by the self-mass term. The final diagram for $\Gamma(q; r)$ does depend on q; the self-mass correction cancels this graph completely only for $q = 0$.

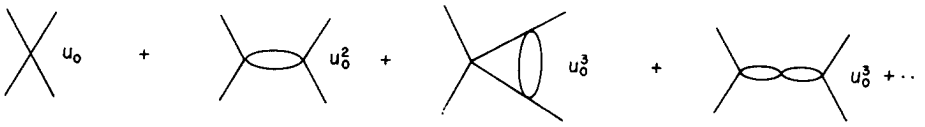

Fig. 8.1. Graphs contributing to u_R.

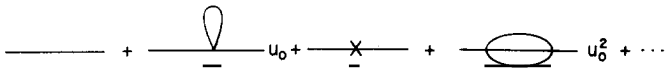

Fig. 8.2. Graphs contributing to $\Gamma(q;r)$. The x refers to the $r_0 - r$ term in (8.8).

As an illustration of the calculation a Feynman graph in 3.99 dimensions consider the u_0^2 contribution to u_R. It gives

$$u_0^2 \int \frac{d^{(d)}q}{[q^2(1+q^2)^2 + r]^2}. \tag{8.18}$$

The angular integration is simple and produces an unimportant constant. Now (8.18) becomes,

$$\text{const. } u_0^2 \int_0^\infty \frac{q^{d-1}\,dq}{[q^2(1+q^2)^2 + r]^2}. \tag{8.19}$$

The ϵ dependence of (8.19) comes from,

$$q^{d-1} = q^{3-\epsilon} = q^3 - \epsilon q^3 \log q + \tfrac{1}{2}\epsilon^2 q^3 \log^2 q - \ldots . \tag{8.20}$$

Consider first the q^3 term contributing to (8.19). We are interested in very small values of r, so the range of the integral $r \ll q^2 \ll 1$ will produce a single $\log r$,

$$\int \frac{q^3\,dq}{[q^2(1+q^2)^2 + r]^2} \sim \int_r^1 \frac{dq^2}{q^2} = -\log r. \tag{8.21}$$

Similarly the $\epsilon q^3 \log q$ term in (8.20) will produce an $\epsilon \log^2 r$ contribution to (8.19). So, we can expect an expansion of (8.19) in ϵ to have the form,

$$u_0^2 \{\ln r + \epsilon[\ln^2 r + \ln r + \text{const.}] + \epsilon^2[\ln^3 r + \ldots]\} \tag{8.22}$$

where a host of numerical constants have been suppressed. The integral in this example was particularly simple. In general we meet nontrivial angular integrations such as,

$$\int q^{d-1} \sin^{d-2}(\theta) \ldots dq\, d\theta \tag{8.23}$$

and possibly even two angular integrations. These are still well-defined for non-integer d [126, 128].

Using the techniques discussed here and exercising more care with factors of 2 and π, one obtains the following expansions:

$$u_R = u_0 + 9\,(u_0^2/4\pi^2)\{\ln r + \tfrac{17}{6} - \tfrac{1}{4}\epsilon \ln r \ldots\} + \ldots \tag{8.24}$$

for small r, and

$$\Gamma(q;0) = \frac{1}{q^2}\left\{1 + \frac{3u_0^2}{8\pi^4}\ln q + \ldots\right\} \tag{8.25}$$

for small q. First observe that (8.25) lacks a u_0 term since no graph enters the calculation in first order. Since $\Gamma(q;0) \sim q^{-2+\eta}$ for q small, η must be of order u_0^2. Now look at u_R. Up to the $\ln r$ term,

$$u_R = u_0\left\{1 + \frac{9u_0}{4\pi^2}\ln r + \ldots\right\} \tag{8.26}$$

which must be proportional to $r^{(\epsilon-2\eta)/(2-\eta)}$. The factor u_0 can be dropped (it is a constant of proportionality); therefore

$$1 + \frac{9u_0}{4\pi^2}\ln r \propto r^{(\epsilon-2\eta)/(2-\eta)} = 1 + [\tfrac{1}{2}\epsilon + O(u_0^2)]\ln r + \ldots \tag{8.27}$$

and we have

$$u_0 = \tfrac{2}{9}\pi^2\epsilon + \text{higher orders.} \tag{8.28}$$

Matching (8.25) with (8.5), we find:

$$\eta = \epsilon^2/54. \tag{8.29}$$

It is interesting to observe that 54 is not a small number, so even choosing $\epsilon = 1$ results in $\eta \approx 0.02$ which is small compared to unity. η has been calculated to order ϵ^3 giving $\eta \approx 0.04$ for $\epsilon = 1$. The best guesses for η from the high temperature expansion give 0.041–0.056 [129, 130], so our rough calculation seems quite reliable. The anomalous dimension of the spin field becomes, for $\epsilon = 1$,

$$d_s = \tfrac{1}{2}(d - 2 + \eta) = 0.5 + 0.025$$

from the high temperature expansion result. It is d_s which has special significance as an anomalous dimension in field theory applications. And it is interesting to note that the correction to the mean field theory value for d_s is small compared to unity.

Table 8.1 lists the results of Feynman graph calculations for critical exponents obtained to date. Table 8.2 lists the comparison of Feynman graph calculations at $\epsilon = 1$ with high temperature series calculations for the three-dimensional Ising model (as summarized in [130]). The differences are too small to be taken seriously.

Table 8.2
Comparison of high temperature results for the 3-dimensional Ising model [130] with the ϵ expansion to order ϵ^2, except to order ϵ^3 for η (table 8.1) for $\epsilon = 1$ and $n = 1$, and Landau theory.

Exponent	ϵ expansion	Ising	Landau
ν	0.626	0.642 ± 0.003	0.5
η	0.037	0.055 ± 0.010	0
γ	1.244	1.250 ± 0.003	1.0
α	0.077	0.125 ± 0.015	0
β	0.340	0.312 ± 0.003	0.5
δ	4.46	5.150 ± 0.02	3

9. Dimension of tensor operators in $4 - \epsilon$ dimensional space-time

In this section quantum field theory in 3.99 space-time dimensions d will be discussed. In particular, the dimensions of tensor operators relevant to deep inelastic scattering will be computed using the ϵ expansion. (For a review of theories of deep inelastic electron scattering and the relevance of tensor operators see [132].) The theory discussed here approaches a free field theory for $d \to 4$, so it is not directly relevant to physics. The importance of the calculation is that it gives an example of anomalous dimensions in a theory which shows scale invariance at short distances. Previously the only known example was the more trivial Thirring model. (For references on the Thirring model see [133].)

The details of the connection between critical phenomena and field theory will be discussed in the following section. Here we shall rely on the obvious connection that the Feynman graphs discussed in previous lectures are similar to the (unrenormalized) Feynman graphs for a ϕ^4 field theory. The actual calculation will be set up in statistical mechanical language; the obvious analogy between correlation functions of spins and vacuum expectation values of fields will allow us to interpret the results of the calculation in terms of anomalous dimensions.

To begin, consider the rules of the Feynman graph approach to the ϵ expansion. The propagator had the cutoff form $[q^2(1+q^2)^2 + r]^{-1}$ and the vertex was $-u_0$. The mass renormalization was done such that the exact propagator at momentum zero was equal to the free propagator at momentum zero, namely r^{-1}. Now consider ϕ^4 field theory. The propagator is $[k^2 - m^2 + i\epsilon]^{-1}$ and the vertex is $i\lambda_0$. One can relate the field theoretic quantities m^2, k^2 and λ_0 to the statistical parameters r, q^2 and u_0. Comparing the propagators at zero momentum, the identification

$$r \leftrightarrow m^2 \tag{9.1}$$

is clear. Considering only small q^2 ($q^2 \ll 1$),

$$q^2 \leftrightarrow -k^2. \tag{9.2}$$

So, positive q^2 corresponds to space-like k. Therefore, the statistical mechanical calculations apply only to vacuum expectation values of fields in the space-like region. It is well-known that the internal momenta in graphs having space-like external legs can be made space-like by a contour rotation. Therefore, the perturbation theory rules of previous lectures are Feynman rules for space-like momenta. The bare coupling after rotation of internal momenta becomes $-\lambda_0$ which should be identified with $-u_0$. m^2 appearing in these rules is not quite the true renormalized mass because the renormalization has been done at zero momentum instead of on the mass shell. Since the cutoff is fixed at 1, one must have small m (i.e. small r) in order that the physical mass be much less than the cutoff. No wave function or coupling constant renormalizations have been done.

The integrations in the statistical mechanical case were of the form $\int d^d q$ where d is the dimension of space. In the field theoretic case d is the number of space and time dimensions. Thus $d = 4 - \epsilon$ means only $3 - \epsilon$ space dimensions in the field theoretic case, whereas it means $4 - \epsilon$ space dimensions for the statistical mechanics case.

In this section we will calculate the dimensions of several composite operators within this framework. Define a vertex in configuration space by,

$$U_{\alpha\beta}(\boldsymbol{q}, -\boldsymbol{q})\, \Gamma^2(q) \equiv \int \exp\{i\boldsymbol{q} \cdot \boldsymbol{x}\}\ \exp\{-i\boldsymbol{q} \cdot \boldsymbol{y}\}\, \langle s(x)\, s(y)\, T_{\alpha\beta}(0)\rangle \tag{9.3}$$

where

$$T_{\alpha\beta}(0) \equiv s(0)\, \nabla_\alpha\, \nabla_\beta\, s(0). \tag{9.4}$$

$\Gamma(q)$, the exact propagator, appears squared in (9.3) in order to remove self-energy insertions on the external legs of the vertex. The operator $T_{\alpha\beta}$ will be denoted graphically as in fig. 9.1. We are particularly interested in the pure tensor piece of the operator $T_{\alpha\beta}$ (as opposed to the scalar part $\Sigma_\alpha T_{\alpha\alpha}$). Choosing $\alpha \neq \beta$ picks out just this piece.

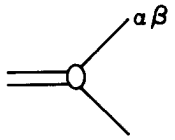

Fig. 9.1. Diagrammatic representation of the composite operator $T_{\alpha\beta}$.

In field theoretic language $T_{\alpha\beta}$ is a linear combination of the stress energy tensor and the total derivative $\nabla_\alpha \nabla_\beta s^2(0)$. The matrix element $U_{\alpha\beta}$ has been cleverly chosen so that the total derivative does not contribute to $U_{\alpha\beta}$. The reason is that the external momentum carried by $T_{\alpha\beta}$ is zero. It can be argued on general grounds that the dimension of the stress energy tensor is d, the dimensionality of the system, for a scale invariant theory. We will see that this is true at least to order ϵ^2. In real life there are other operators which carry internal quantum numbers such as isospin which need not have dimension d. To illustrate this, internal quantum numbers will be introduced by giving s an internal index. Call the new fields s_i. The theory will be constructed to be invariant under rotations among these fields. The interaction term in the new Hamiltonian becomes $\Sigma_{i,j} u_0 s_i^2 s_j^2$. There are two possible tensors:

$$T_{\alpha\beta}(0) = \sum_j s_j(0)\ \nabla_\alpha \nabla_\beta s_j(0) \tag{9.5a}$$

and,

$$T_{\alpha\beta ij}(0) = s_i(0)\nabla_\alpha \nabla_\beta s_j(0) \quad (i \neq j). \tag{9.5b}$$

$T_{\alpha\beta ij}(0)$ is a tensor with respect to the internal rotational symmetry as well as to spatial rotations. Corresponding vertices $U_{\alpha\beta}(q, -q)$ are defined by substituting $\langle s_j(x) s_j(y) T_{\alpha\beta}(0)\rangle$ or $\langle s_i(x) s_j(y) T_{\alpha\beta ij}(0)\rangle$ into (9.3).

The analysis of these tensors will be done for small r and small external momenta q. In order to discuss the case $q \to 0$, a kinematic factor $q_\alpha q_\beta$ must be removed from $U_{\alpha\beta}$ explicitly. We will calculate the scalar functions multiplying $q_\alpha q_\beta$. To begin, recall several results from the last lecture. Scaling arguments proved that the exact propagator at zero momentum scales as a power of the correlation length, ξ^{d-2d_s}. Similarly, the four point function scales as $\xi^{3d-4d_s}/\xi^{4d-8d_s}$ when its external momenta are all set to zero. These results are easily understood. In the case of the propagator, one is considering the quantity $\int d^d x \langle s(x) s(0)\rangle$. The important values of x which contribute to the integral are $x \sim \xi$. Therefore, the d-dimensional integral contributes ξ^d while each spin field contributes a factor ξ^{-d_s}. Similar arguments can be made for the quantity $U_{\alpha\beta}(q, -q)$ and $U_{\alpha\beta ij}(q, -q)$ when q is very small. If d_T is the anomalous dimension of the operator $T_{\alpha\beta}$, then

$$U_{\alpha\beta}(q, -q) \sim \frac{\xi^{2d-2d_s-d_T}}{\xi^{2d-4d_s}} F_{\alpha\beta}(q\xi). \tag{9.6}$$

As argued above, $U_{\alpha\beta}(q, -q)$ is proportional to $q_\alpha q_\beta$. Therefore, for small q, $F_{\alpha\beta}$ is proportional to $q_\alpha \xi\, q_\beta \xi$. Then (9.6) becomes,

$$U_{\alpha\beta}(q, -q) \sim \xi^{2+2d_s-d_T} q_\alpha q_\beta \quad (q\xi \ll 1). \tag{9.7}$$

Eq. (9.7) can be written in terms of r and η using the definitions $d_s = \frac{1}{2}(d - 2 + \eta)$ and $r^{-1} \sim \xi^{d-2d_s}$. Then (9.7) becomes,

$$U_{\alpha\beta}(\boldsymbol{q}, -\boldsymbol{q}) \sim r^{(d_T - d - \eta)/(2-\eta)} q_\alpha q_\beta \quad (q\xi \ll 1). \tag{9.8}$$

The critical exponent η can be calculated as a series in ϵ by using the matching procedure described in section 8. The only added ingredient is the presence of internal symmetry which increases the number of graphs which must be considered. The appropriate generalizations of (8.24) and (8.25) turn out to be,

$$r^{\epsilon/2} \sim u_0 \{1 + 4(n+8)\frac{u_0}{16\pi^2} \ln r + \ldots\}, \qquad q^\eta \sim 1 + 32(n+2)\frac{u_0^2}{(16\pi^2)^2} \ln q + \ldots, \tag{9.9}$$

where n is the number of possible values for the internal symmetry label i. And finally, u_0 and η can be expressed in terms of ϵ,

$$u_0/16\pi^2 = \epsilon/8(n+8), \qquad \eta = 32(n+2)\, u_0^2/(16\pi^2)^2. \tag{9.10}$$

Setting $n = 1$ reproduces the results of section 8.

The same type of analysis will now be used to determine dimensions of tensor operators. Consider the diagrams contributing to the vertex functions $U_{\alpha\beta}(\boldsymbol{q}, -\boldsymbol{q})$ and $U_{\alpha\beta ij}(\boldsymbol{q}, -\boldsymbol{q})$. To order zero in u_0 the diagrams are those of fig. 9.2. The possible diagrams in order u_0 and u_0^2 are shown

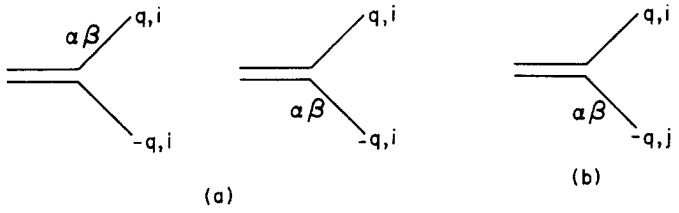

Fig. 9.2. (a) Zeroth order graphs contributing to $U_{\alpha\beta}$. (b) Zeroth order graph contributing to $U_{\alpha\beta,\, ij}$.

in figs. 9.3 and 9.4. However, the diagrams of figs. 9.3 and 9.4a vanish. The reason for this is that q does not appear in the loop integration where the indices α and β appear. Therefore, there is no possible way for a factor $q_\alpha q_\beta$ to emerge from the calculation. Hence the graph must vanish identically. The calculation of the graph shown in fig. 9.4b consists of two parts. First is the calculation of the integrals and second is counting the number of ways the graph can be constructed.

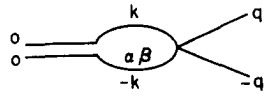

Fig. 9.3. Potential contribution to $U_{\alpha\beta}$ of order u_0. This diagram vanishes.

One way to calculate the basic graph in fig. 9.4b is to do the internal p integration first,

$$\Sigma[(\boldsymbol{k}+\boldsymbol{q})^2] = \int \frac{1}{[p^2(1+p^2)^2 + r]} \frac{1}{[(\boldsymbol{p}+\boldsymbol{k}+\boldsymbol{q})^2[1 + (\boldsymbol{p}+\boldsymbol{k}+\boldsymbol{q})^2]^2 + r]} \, \mathrm{d}^d p. \tag{9.11}$$

The k integration then reads,

$$u_0^2 \int k_\alpha k_\beta \frac{1}{[k^2(1+k^2)^2+r]^2} \Sigma[(k+q)^2]\, \mathrm{d}^d k. \tag{9.12}$$

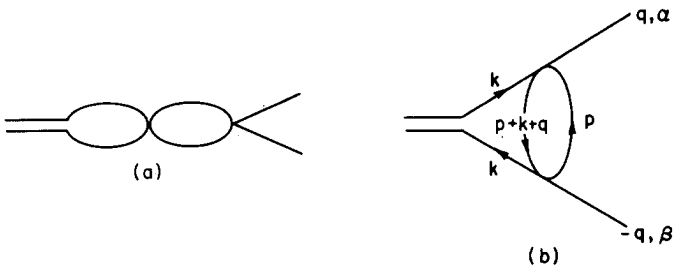

Fig. 9.4. Two graphs of order u_0^2 contributing to $U_{\alpha\beta}$. Graph (a) vanishes identically but (b) does not.

Since our goal is to calculate $U_{\alpha\beta}$ and $U_{\alpha\beta ij}$ to $O(u_0^2)$, the $d = 4 - \epsilon$ dimensional integrals can be replaced by familiar four dimensional integrals. This is true since (9.12) is $O(u_0^2)$ and the ϵ dependence of the integrals will introduce only higher order corrections. Eq. (9.12) is not difficult to evaluate. The contribution to $U_{\alpha\beta}$ in the limit $q \to 0$ is,

$$-\tfrac{1}{6} q_\alpha q_\beta \frac{u_0^2}{(16\pi^2)^2} \ln r. \tag{9.13}$$

To do the counting one must look at the different ways the indices i and j can run through the graph. Fig. 9.5a, b shows two alternatives for $U_{\alpha\beta}$ while fig. 9.6a, b shows two alternatives for $U_{\alpha\beta ij}$. Associated with fig. 9.5a is a counting factor $4 \times 4 \times 2n$. The factor n comes from the free sum over indices of the internal loop. In fact, any graph with an internal loop of indices contributes a term proportional to n. Furthermore, there are 4 ways the upper four point vertex could

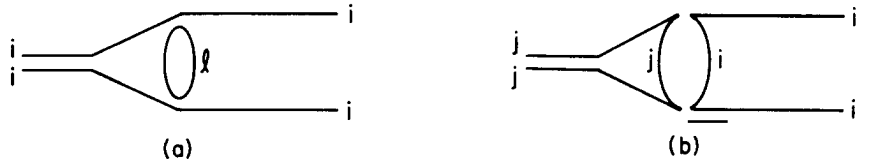

Fig. 9.5. Graphs (a) and (b) show two distinct ways for indices to circulate through a graph contributing to $U_{\alpha\beta}$.

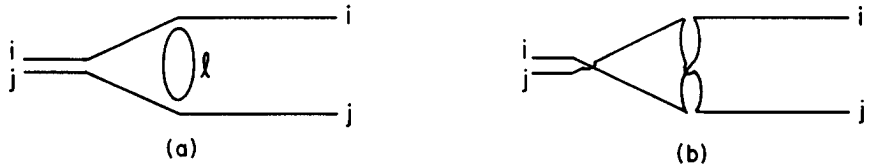

Fig. 9.6. Graphs (a) and (b) depict two alternate ways for indices to appear in graphs contributing to $U_{\alpha\beta,\, ij}$.

be constructed. A similar factor of 4 comes from the bottom vertex. Finally, there is a factor of 2 from the ways of making the loop after the two vertices are determined. Similarly, the counting factor for fig. 9.5b is $64n$. The total for the two figures is $96n$. There are also diagrams without loops (of the sort in fig. 9.6b) which contribute to $U_{\alpha\beta}$. Taking these into account, the total counting factor is $96(n+2)$. The counting problem for $U_{\alpha\beta,\,ij}$ involves different combinatorics. Fig. 9.6a contributes $32n$ (the calculation is identical to fig. 9.5a). However, this is the only graph with a loop contributing to $U_{\alpha\beta,\,ij}$. Fig. 9.6b shows one of the non-loop diagrams. Including all the possibilities produces a counting factor of $32(n+6)$ for $U_{\alpha\beta,\,ij}$.

Collecting the integrals and the counting factors gives,

$$U_{\alpha\beta} \sim 2q_\alpha q_\beta \left[1 - \frac{96(n+2)}{6} \frac{u_0^2}{(16\pi^2)^2} \ln r\right]. \tag{9.14}$$

The overall factor of 2 in the $O(u_0^2)$ term comes from the fact that the indices α and β could appear on either the upper or lower lines in fig. 9.5a, b etc. Similarly,

$$U_{\alpha\beta,\,ij} \sim q_\alpha q_\beta \left[1 - \frac{32(n+6)}{6} \frac{u_0^2}{(16\pi^2)^2} \ln r\right]. \tag{9.15}$$

Eq. (9.14) is to be compared with the scaling result,

$$q_\alpha q_\beta r^{(d_T - d - \eta)/(2-\eta)}. \tag{9.16}$$

Since η is $O(u_0^2)$, it can be neglected in the denominator of the exponent. Then, comparing (9.16) with (9.14) gives,

$$d_T = d + \eta - \tfrac{96}{3}(n+2)\, u_0^2/(16\pi^2)^2. \tag{9.17}$$

Using the previous result (9.10) for η it is apparent that the second and third terms cancel in (9.17) leaving

$$d_T = d \tag{9.18}$$

as claimed. To determine the anomalous dimension $d(T_{ij})$ for $T_{\alpha\beta ij}$ compare (9.15) with (9.16). This gives

$$d(T_{ij}) = d + \eta - \tfrac{32}{3}(n+6)\, u_0^2/(16\pi^2)^2. \tag{9.19}$$

Using (9.10) this becomes,

$$d(T_{ij}) = d + \frac{n}{3(n+8)^2}\, \epsilon^2. \tag{9.20}$$

Choosing $n = 3$ (so the s_i form an isospin triplet) gives,

$$d(T_{ij}) = d + \epsilon^2/121. \tag{9.21}$$

Since ϵ is at most of order unity, the deviation of $d(T_{ij})$ from d is remarkably small. Of course, higher order corrections to $d(T_{ij})$ will change the details of $d(T_{ij}) - d$, but the semi-quantitative result that $d(T_{ij}) - d$ is small is expected to still hold. The reason the anomaly is small is that the graph of order u_0 (fig. 9.3) is identically zero.

The relevance of this calculation is, of course, a matter of debate. However, it does present an example of a scale invariant theory in which one can actually calculate anomalous dimensions. It

may be that the anomalous dimensions that actually govern the behavior of deep inelastic scattering [132, 134] are also very close to canonical. (For recent perturbation theoretic analyses of deep inelastic electron scattering see refs. [135, 136].)

Since u_0 is of order ϵ, it is clear that $u_0 \to 0$ for $\epsilon \to 0$ which means there is no interaction when $d \to 4$. So this calculation cannot be applied directly to real elementary particle physics. The hope is that what one sees here for dimension $d < 4$ (say $d = 3$) is an indication of what a more realistic theory will show for $d = 4$.

Anomalous dimensions have been computed for the mth rank tensor operators T:

$$T_{\alpha_1 \ldots \alpha_m}(0) = \sum_j s_j(0) \nabla_{\alpha_1} \ldots \nabla_{\alpha_m} s_j(0) \tag{9.22}$$

$$T_{\alpha_1 \ldots \alpha_m ij}(0) = s_i(0) \nabla_{\alpha_1} \ldots \nabla_{\alpha_m} s_j(0) \tag{9.23}$$

for even m. These will be denoted $T_{(m)}$ and $T_{(m)ij}$, respectively. The results, to order ϵ^2, are

$$d(T_{(m)}) = \frac{(n+2)}{2(n+8)^2} \epsilon^2 \left\{1 - \frac{6}{m(m+1)}\right\}, \tag{9.24}$$

$$d(T_{(m)ij}) = \frac{(n+2)}{2(n+8)^2} \epsilon^2 \left\{1 - \frac{2(n+6)}{(n+2)m(m+1)}\right\}. \tag{9.25}$$

10. Connection between statistical mechanics and field theory

In section 9 a brief argument was given that the Feynman diagrams for the ϕ^4 interaction (converted to imaginary times, i.e., a Euclidean metric) are identical to the diagrams for the partition function of earlier lectures. To be precise, the vacuum expectation value of n fields ϕ is proportional to the n-spin correlation function. In this lecture the same identity will be established at a more fundamental level and with greater care. Greater care means, for example, introducing a specific (and noncovariant) cutoff procedure for the ϕ^4 theory, and not making use of the Feynman expansion.

The connection between critical phenomena and field theory has been recognized and exploited in diagrammatic form by Gribov and Migdal [137, 138] and Polyakov [139–141]; a detailed comparison of the Feynman path integral to the partition function is given by Moore [142]. The connection is also discussed in a thesis by Suri [143]. The emphasis in this section is on the transfer matrix (as in [143]).

Quantum theory on a lattice will be a crucial part of the discussion. In this theory one has discrete operators ϕ_n instead of a quantum field operator $\phi(x)$. Simple Hamiltonians on a lattice have the form of coupled harmonic or anharmonic oscillators (see below). A locality requirement will be imposed: the Hamiltonian H should be a sum,

$$H = \sum_m H_m \tag{10.1}$$

where H_m depends only on operators ϕ_n near the site m. Quantum field theory reduces to a lattice theory by use of a suitable cutoff method; lattice theories will be defined from statistical mechanics using the "transfer matrix" formalism (see, e.g., [144]).

The cutoff procedures are likely to distort the original ϕ^4 theory even if the original theory makes sense. In perturbation theory the effects of a large but finite cutoff can be removed by standard perturbation theoretic renormalization techniques. A nonperturbative renormalization

theory will be described in section 12. At the end of this section there will be some preliminary remarks establishing the equivalence of an infinite cutoff in the field theory to an infinite correlation length in the statistical mechanics.

The lattice theories are not Lorentz invariant (or Euclidean invariant for imaginary times). Lorentz invariance has to be recovered in a continuum limit. The nature of this limit is determined by the renormalization process. In perturbation theory Lorentz invariance is restored. (For example, the graphs in sections 13 and 9 are all Euclidean invariant.) The conditions needed to give Euclidean invariance in general are discussed in section 12.

There is no trivial connection between the Hamiltonian which defines the statistical mechanics and the Hamiltonian of the corresponding field theory defined in this section. If the statistical mechanical Hamiltonian is defined in a space of dimension d, the field theoretic Hamiltonian acts in a space of dimension $d-1$. Statistical mechanical interactions are denoted by $\mathcal{H}$, field theoretic Hamiltonians by H. The statistical mechanical interaction $\mathcal{H}$ is more closely related to the Lagrangian of the field theory (see below and [142]).

The transfer matrix formalism has been described elsewhere [144], but will be reviewed here. (It was used in [144] as a basis for a transparent derivation of Onsager's solution of the two-dimensional Ising model.) Consider a statistical mechanical model on a discrete lattice. The partition function is defined to be,

$$Z = \prod_m \left(\int ds_m \exp\left\{ -\frac{b}{2} s_m^2 - u_0 s_m^4 \right\} \right) \exp\left\{ K \sum_n \sum_{\hat{i}} s_n s_{n+\hat{i}} \right\} \tag{10.2}$$

where the notation is borrowed from sections 3 and 4. (The sum over $\hat{i}$ is a sum over the d axes; $\hat{i}$ is a unit vector along the axis i.) The partition function (10.2) involves integrations over the spins at each lattice site. Inside the exponentials there are terms associated with individual sites and terms which couple nearest neighbor sites. We will now show by construction that Z can be written in the form,

$$Z = \operatorname{tr} V^N \tag{10.3}$$

where N is the extent of the lattice (assumed finite for now) and V is the transfer-matrix, which will be defined below. To begin, one must organize the integrations in (10.2). For a plane lattice the spin variables are integrated out row by row (see fig. 10.1). (For a lattice of dimension d, a

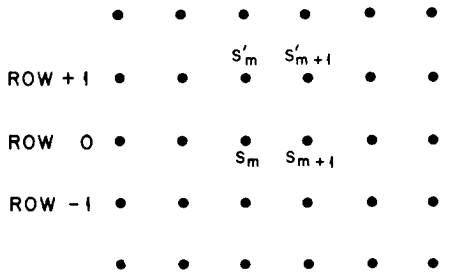

Fig. 10.1. A two-dimensional lattice of spins. Spins within each row are labelled with a subscript.

$d-1$ dimensional sublattice substitutes for a row.) Schematically, Z will be written in the form,

$$Z = \ldots \int_{\text{row 1}} \ldots \int_{\text{row 0}} \ldots \int_{\text{row } -1} \ldots \tag{10.4}$$

where the rows have been labelled in fig. 10.1. Since the interaction in (10.2) involves only nearest neighbors, the partition function can be written more explicitly in the form,

$$Z = \ldots \int_{\text{row 1}} \exp\{\mathcal{H}[\text{row 2, row 1}]\} \int_{\text{row 0}} \exp\{\mathcal{H}[\text{row 1, row 0}]\} \int \ldots \tag{10.5}$$

There is a degree of ambiguity in the meaning of $\mathcal{H}$ in (10.5) since terms in (10.2) which involve only row 1, say, could appear in either $\mathcal{H}$[row 2, row 1] or $\mathcal{H}$[row 1, row 0]. By convention such terms will be divided equally between $\mathcal{H}$[row 2, row 1] and $\mathcal{H}$[row 1, row 0]. Clearly all the $\mathcal{H}$'s in (10.5) are identical in form and differ only by a translation. It is convenient to label lattice sites within a row (sublattice) by a $d-1$ dimensional vector $\boldsymbol{m}$. (For a two dimensional lattice $\boldsymbol{m}$ is a scalar.) Defining $s_{\boldsymbol{m}}$ ($s'_{\boldsymbol{m}}$) to be the spins in row 0 (row 1) (fig. 10.1), it follows from (10.2) that,

$$\mathcal{H}[\text{row 1, row 0}] = -\tfrac{1}{4} b \sum_{\boldsymbol{m}} (s_{\boldsymbol{m}}^2 + s_{\boldsymbol{m}}'^2) - \tfrac{1}{2} u_0 \sum_{\boldsymbol{m}} (s_{\boldsymbol{m}}^4 + s_{\boldsymbol{m}}'^4) + \tfrac{1}{2} K \sum_{\boldsymbol{m},\hat{\imath}} (s_{\boldsymbol{m}} s_{\boldsymbol{m}+\hat{\imath}} + s'_{\boldsymbol{m}} s'_{\boldsymbol{m}+\hat{\imath}}) + K \sum_{\boldsymbol{m}} s_{\boldsymbol{m}} s'_{\boldsymbol{m}}, \tag{10.6}$$

where the third term includes interactions among the lattice sites in rows 0 and 1. The fourth term gives the interactions between row 0 and row 1.

Let s denote the set of spins $\{s_{\boldsymbol{m}}\}$ in a particular row. Then $\mathcal{H}$ is a function of s and s'. Define

$$V(s', s) = \exp\{\mathcal{H}(s', s)\}. \tag{10.7}$$

The function $V(s', s)$ defines the elements of a matrix; matrix multiplication involves integration over all the spins $s_{\boldsymbol{m}}$ in s. One can now set up an operator-wave function formalism. Wave functions are functions $\psi(s)$; the operator V acting on ψ gives a wave function $\psi'(s')$:

$$\psi'(s') = \int_s V(s', s)\psi(s) \tag{10.8}$$

where $\int_s$ means $\Pi_{\boldsymbol{m}} \int_{-\infty}^{\infty} \mathrm{d}s_{\boldsymbol{m}}$.

On a lattice of finite extent it is convenient to impose periodic boundary conditions, which means only that, in addition to coupling row n to row $n+1$ ($1 \leqslant n \leqslant N-1$, say), one couples row N to row 1. In this case, the partition function is $\mathrm{Tr}(V^N)$ as promised.

The space of functions $\psi(s)$ defines a Hilbert space (provided each row is of finite extent; otherwise one has an inseparable Hilbert space, which makes the rigorously-minded croak). The operator V is an Hermitian operator since $\mathcal{H}(s', s)$ is real and symmetric. Therefore, V could be chosen to be a Hamiltonian of some quantum mechanical system. However, an acceptable Hamiltonian must satisfy the additional requirement of locality. V is the exponential of sums over the lattice sites. To obtain a quantity which is additive over distant lattice sites, the Hamiltonian of the field theory will be chosen to be,

$$H = -\frac{1}{\tau} \ln V \tag{10.9}$$

where τ is an arbitrary normalization constant. When K in (10.6) is zero it is easy to show that $\ln V$ is actually a sum over lattice sites (in this case V is the direct product of independent operators, one for each site). This can also be verified in perturbation theory in K [143]. Otherwise proving the locality of $\ln V$ is a nontrivial problem which will not be discussed further.

To put the definition (10.9) of H into perspective, a long digression will be made on the definition of the Hamiltonian and the nature of the operator $\exp(-H\tau)$ in quantum mechanics and quantum field theory.

In the ordinary formulation of quantum mechanics there are two types of restrictions on the Hamiltonian. First there are general restrictions on H; it must be Hermitian, it must obey specific symmetries in specific cases, etc. In field theory H must also be at least macroscopically local. The second type of restriction is a requirement of simplicity. The Coulomb Hamiltonian can be written explicitly in terms of just two parameters, the electron charge and mass. This contrasts with the set of eigenvalues of the helium atom, for example, which are extremely complicated and cannot be written down in closed form. The Hamiltonian would be a much less useful concept if there were no simple way to define it.

Even in cases like nuclei where the exact Hamiltonian is not known, the tendency is to invent simple model Hamiltonians (square well or harmonic oscillator potential models, for example) which reproduce qualitative features of the physics.

In the present case the requirement of simplicity has been ignored. The operator V has been rather simply defined, but $\ln V$ is more complicated. For example, one cannot write matrix elements $\langle s'|\ln V|s\rangle$ in closed form. The field theories one normally investigates (e.g., quantum electrodynamics or the ϕ^4 interaction of a scalar field) have simple Hamiltonians. At least this is the case if canonical field theory is to be believed. There are three arguments for being interested in theories with less simple Hamiltonians such as the one defined here. First is that one may be interested in studying Hamiltonians of any kind that satisfy general principles such as locality. Then the advantage of considering $\ln V$ as a Hamiltonian is that all the techniques of statistical mechanics are available to solve it. It will become evident in section 14 that statistical mechanical methods are more powerful for strongly interacting (scalar) theories than any field theoretic methods. The second argument is based on the universality principle which will be discussed in section 12. What will be found in section 12 is that large classes of interactions are expected to give the same renormalized field theory, so the precise form of the initial interaction (simple or complex) is not very important. Finally, there are cutoff procedures which result in the form (10.9) for H; these cutoff procedures will be explained later in this section.

Now consider standard quantum mechanics and the operator $\exp(-H\tau)$. If one defines a state vector $\psi(\tau)$ by,

$$|\psi(\tau)\rangle = \exp(-H\tau)|\psi\rangle \tag{10.10}$$

where ψ is τ independent, then $\psi(\tau)$ satisfies the differential equation,

$$\frac{\partial}{\partial\tau}|\psi(\tau)\rangle = -H|\psi(\tau)\rangle. \tag{10.11}$$

This is just the Schrödinger equation for imaginary time. Suppose that the state $|\psi(\tau)\rangle$ has a wave function $\psi(x, \tau)$. Then the solution to (10.11) can be written in the form,

$$\psi(x, \tau) = \int_{-\infty}^{\infty} G(x, \tau, x', 0)\psi(x', 0)\,dx' \tag{10.12}$$

where G is a Green's function: in abstract form,

$$G(x, \tau, x', 0) = \langle x|\exp(-H\tau)|x'\rangle. \tag{10.13}$$

As discovered by Dirac and Feynman (see [145] or [146]), the Green's function has a simple form for small τ. One can write formally,

$$G(x, \tau, x', 0) = \exp\left\{\int_0^\tau \bar{L}[x, \dot{x}]\, d\tau\right\} \tag{10.14}$$

where $\bar{L}$ is the Lagrangian of the non-relativistic system with the time t replaced by $-i\tau$. For a simple time independent potential V, the Lagrangian for imaginary time reads,

$$\bar{L} = -\tfrac{1}{2} m\dot{x}^2 - V(x) \tag{10.15}$$

(where $\dot{x}$ is $dx/d\tau$). For small τ, $\dot{x}$ is approximately,

$$\dot{x} = \{x(\tau) - x(0)\}/\tau. \tag{10.16}$$

Then (10.14) becomes for small τ,

$$G(x, \tau; x', 0) \approx \exp\left\{-\frac{1}{2} m \frac{(x-x')^2}{\tau} - \tau\left[\frac{V(x) + V(x')}{2}\right]\right\}. \tag{10.17}$$

This equation provides the definition in practice of eq. (10.14). The proof that (10.17) gives the Green's function for small τ is given elsewhere [145, 146].

Eq. (10.17) shows that $\exp(-H\tau)$ is a simple operator for small τ. If τ is not small one can write,

$$\exp(-H\tau) \approx \exp(-H\,\tau/l)\exp(-H\,\tau/l)\ldots\exp(-H\,\tau/l) \qquad (l \text{ factors}) \tag{10.18}$$

where l is so large that τ/l is small. The operators $\exp(-H\,\tau/l)$ are simple, and $\exp(-H\tau)$ is built from them by iteration. If one computes the multiple product by matrix multiplication, one has

$$\langle x|\exp(-H\tau)|x'\rangle = \int_{x_1} \ldots \int_{x_l} \langle x|\exp(-H\tau/l)|x_1\rangle\langle x_1|\exp(-H\tau/l)|x_2\rangle \ldots \langle x_l|\exp(-H\tau/l)|x'\rangle. \tag{10.19}$$

This multiple integral can be written more explicitly using eq. (10.17). In the limit $l \to \infty$ it defines the Feynman path integral [145, 146].

In quantum field theory there is also a path integral formalism; one must introduce a cutoff to make it well defined. One can introduce the cutoff in such a way that the Hamiltonian continues to have a simple form and likewise for $\exp(-H\tau)$ for infinitesimal τ. Alternatively, the cutoff can be introduced so that only $\exp(-H\tau)$ for a given finite value of τ is simple. This will be explained below. Neither cutoff method is Lorentz invariant; there are of course Lorentz invariant cutoff methods too. (A Euclidean invariant cutoff was used in sections 8 and 9.)

Consider the Lagrangian for a scalar field theory (in $d-1$ space, 1 time dimension):

$$L = \int d^{d-1}x\{\tfrac{1}{2}(\partial\phi/\partial t)^2 - \tfrac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}\mu_0^2\phi^2 - \lambda_0\phi^4\} \tag{10.20}$$

where $\phi(x)$ is the scalar field. As the first cutoff method, replace continuous space by a lattice of small but non-zero spacing a. One uses a finite difference approximation for the gradient and replaces the integral by a sum. Let ϕ_m be the value of $\phi(x)$ at lattice site m ($x = ma$). Then,

$$L = a^{d-1}\sum_m \left\{\tfrac{1}{2}(\partial\phi_m/\partial t)^2 - \frac{1}{2a^2}\sum_{\hat{i}}(\phi_{m+\hat{i}} - \phi_m)^2 - \tfrac{1}{2}\mu_0^2\phi_m^2 - \lambda_0\phi_m^4\right\}. \tag{10.21}$$

Eq. (10.21) is the Lagrangian for a set of coupled anharmonic oscillators, with each ϕ_m being the displacement of an oscillator. The mass parameter of each oscillator is a^{d-1}. If P_m is the momentum operator for the oscillator at m, the Hamiltonian is,

$$H = \sum_m \left\{ \frac{1}{2a^{d-1}} P_m^2 + \tfrac{1}{2} a^{d-3} \sum_{\hat{i}} (\phi_{m+\hat{i}} - \phi_m)^2 + \tfrac{1}{2} \mu_0^2 a^{d-1} \phi_m^2 + \lambda_0 a^{d-1} \phi_m^4 \right\}. \tag{10.22}$$

The operator $\exp(-H\tau)$ for sufficiently small τ is (by analogy with the non-relativistic formula (10.17)),

$$\langle \phi' | e^{-H\tau} | \phi \rangle \approx \exp \sum_m \left\{ -\frac{1}{2\tau} a^{d-1} (\phi'_m - \phi_m)^2 - \tfrac{1}{4} \tau a^{d-3} \sum_{\hat{i}} [(\phi_{m+\hat{i}} - \phi_m)^2 + (\phi'_{m+\hat{i}} - \phi'_m)^2] \right.$$
$$\left. - \tfrac{1}{4} \tau a^{d-1} \mu_0^2 (\phi_m^2 + \phi_m'^2) - \tfrac{1}{2} \tau a^{d-1} \lambda_0 (\phi_m^4 + \phi_m'^4) \right\}. \tag{10.23}$$

How small must τ be for this formula to be valid? To obtain a rough estimate, consider the Hamiltonian (10.22) and neglect the ϕ^4 terms and the terms $\phi_m \phi_{m+\hat{i}}$ which couple different lattice sites. Then the Hamiltonian becomes,

$$H \sim \sum_m \left\{ \frac{1}{2a^{d-1}} P_m^2 + [\tfrac{1}{2} \mu_0^2 a^{d-1} + (d-1) a^{d-3}] \phi_m^2 \right\} \tag{10.24}$$

which is a sum of independent harmonic oscillators with energy level spacing,

$$\omega = \sqrt{\mu_0^2 + 2(d-1)a^{-2}}. \tag{10.25}$$

Since ω is the only energy scale in H, one expects (10.23) to be a good approximation when $\omega\tau \ll 1$, which for small enough lattice spacing a, is just $\tau \ll a$.

The expression (10.23) is rather similar to the formulae (10.6) and (10.7) for $\langle s' | V | s \rangle$. To make an explicit comparison let s_m be the same as ϕ_m except for a scale factor,

$$s_m = \zeta \phi_m. \tag{10.26}$$

Then eq. (10.23) becomes

$$\langle s' | \exp(-H\tau) | s \rangle = \exp \sum_m \{ -\tfrac{1}{4} b (s_m^2 + s_m'^2) - \tfrac{1}{2} u_0 (s_m^4 + s_m'^4) + \tfrac{1}{2} K_1 \sum_i (s_m s_{m+\hat{i}} + s'_m s'_{m+\hat{i}})$$
$$+ K_2 s_m s'_m \} \tag{10.27}$$

where

$$b = \frac{a^d}{\zeta^2} \left\{ \frac{2}{\tau a} + \frac{2(d-1)\tau}{a^3} + \frac{\mu_0^2 \tau}{a} \right\}$$
$$u_0 = \lambda_0 a^{d-1} \tau / \zeta^4, \qquad K_1 = a^{d-3} \tau / \zeta^2, \qquad K_2 = a^{d-1} / \tau \zeta^2. \tag{10.28}$$

The requirement that τ be much less than a means that

$$K_1 \ll K_2. \tag{10.29}$$

Nevertheless, one can still define $\exp(-H\tau) = V$ to be the transfer matrix of a statistical mechanical system; the resulting statistical mechanical interaction is anisotropic in space, having coupling K_2 along one axis, K_1 along all other axes.

Thus, if one is willing to accept anisotropic statistical mechanical interactions, one can have both a simple cutoff field theoretic Hamiltonian and the relation $\tau H = -\ln V$ giving H in terms of a transfer matrix.

It is simpler in practice to deal with isotropic statistical mechanical interactions, which one obtains by putting $\tau = a$ so that $K_1 = K_2$. Now eqs. (10.22) and (10.23) are incompatible. If one is willing to give up the simple explicit form (10.22) for H one can use eq. (10.23) to define H; this means that $H = -\tau^{-1} \ln V$ as proposed earlier (with $K = K_1 = K_2$).

In summary one can take the formal field theoretic expression (analogous to (10.14)),

$$\langle \phi' | \exp(-H\tau) | \phi \rangle = \exp[-\int_0^\tau \int d^{d-1}x \{\tfrac{1}{2}(\partial\phi/\partial\tau)^2 + \tfrac{1}{2}(\nabla\phi)^2 + \tfrac{1}{2}\mu_0^2\phi^2 + \lambda_0\phi^4\} \tag{10.30}$$

and cut this off by introducing a spatial lattice of spacing a. One then has a spatial cutoff length a and a time cutoff τ. If $\tau = a$, one obtains an isotropic transfer matrix and a non-simple H. If $\tau \ll a$, one obtains an anisotropic transfer matrix and a simple H. Eqs. (10.28) give the relation between the field theoretic parameters μ_0 and λ_0, the cutoff a and renormalization parameter ζ, and the statistical mechanical parameters b, u_0, K_1 and K_2. From now on, only the case $\tau = a (K_1 = K_2 = K)$ will be discussed.

The next topic is to establish the connection between spin-spin correlation functions and the vacuum expectation values of the cutoff quantum field theory. The propagator will be discussed explicitly, but the results are valid for any n-field expectation value. The relations $H = -\tau^{-1} \ln V$ and $\zeta\phi_m = s_m$ will be assumed. The standard definition of the propagator, in coordinate space, is

$$D(\boldsymbol{x}, t) = \langle \Omega | T\phi(\boldsymbol{x}, t)\phi(\boldsymbol{0}, 0) | \Omega \rangle \tag{10.31}$$

where $|\Omega\rangle$ is the ground state (vacuum state) and T is the time-ordering operator. The field $\phi(\boldsymbol{x}, t)$ is the Heisenberg field operator

$$\phi(\boldsymbol{x}, t) = \exp(iHt)\phi(\boldsymbol{x}) \exp(-iHt). \tag{10.32}$$

Now replace the continuous variable $\boldsymbol{x}$ by the lattice variable $\boldsymbol{m}$. The result is a lattice propagator $D_{\boldsymbol{m}}(t)$ defined as,

$$D_{\boldsymbol{m}}(t) = \langle \Omega | T\phi_{\boldsymbol{m}}(t)\phi_{\boldsymbol{0}}(0) | \Omega \rangle. \tag{10.33}$$

Using the lattice analogue of (10.31), one has, for $t > 0$,

$$D_{\boldsymbol{m}}(t) = \langle \Omega | \phi_{\boldsymbol{m}} \exp(-iHt)\phi_{\boldsymbol{0}} | \Omega \rangle \exp(iE_0 t) \tag{10.34}$$

where E_0 is the ground state energy of H.

We need the formula for the spin-spin correlation function in terms of the transfer matrix V (see, e.g., [144]). A point on the statistical mechanical lattice is labelled by a row index n and a lattice variable $\boldsymbol{m}$ within a row. A general lattice spin is $s_{n,\boldsymbol{m}}$. The spin-spin correlation function is

$$\Gamma_{n,\boldsymbol{m}} = Z^{-1} \left\{ \ldots \int_{\text{row } n} \ldots \int_{\text{row } 0} \ldots \right\} s_{n,\boldsymbol{m}} s_{0,0} \exp \mathcal{H}. \tag{10.35}$$

(This is for $n > 0$; for $n < 0$ the integrations over row n and row 0 appear in reverse order.) Suppose there are $2N + 1$ rows altogether numbered $-N$ to N, with periodic boundary conditions (and $N \geqslant n$). Then it is easy to see that for $n \geqslant 0$,

$$\Gamma_{n,\boldsymbol{m}} = \mathrm{Tr}(V^{N-n} s_{\boldsymbol{m}} V^n s_0 V^{N+1}), \tag{10.36}$$

(note that one of the factors V couples site N to site $-N$) where $s_{\boldsymbol{m}}$ is now the Schrödinger

operator s_m acting in the Hilbert space of wave functions $\psi(s)$ defined earlier. The trace can be rewritten,

$$\Gamma_{n,m} = \mathrm{Tr}(V^{2N+1-n} s_m V^n s_0)/\mathrm{Tr}(V^{2N+1}). \tag{10.37}$$

Assume that the Hermitian operator V has a unique largest eigenvalue. (Statistical mechanical considerations suggest that this is true for $T > T_c$ ($K < K_c$) and not true for $K > K_c$: see, e.g., [144]). The largest eigenvalue of V is $\exp(-E_0\tau)$ where E_0 is the lowest energy eigenvalue of $H = -(\ln V)/\tau$. The corresponding eigenstate of V is the ground state $|\Omega\rangle$ of H. For large N the operator V^N is

$$V^N \approx |\Omega\rangle \exp(-NE_0\tau) \langle\Omega|; \tag{10.38}$$

the error goes to zero exponentially in N as $N \to \infty$. Using this result, one obtains (for infinite N)

$$\Gamma_{n,m} = \langle\Omega|s_m \exp(-nH\tau)\, s_0|\Omega\rangle \exp(nE_0\tau). \tag{10.39}$$

One can now relate $D_m(t)$ to $\Gamma_{n,m}$. The result is (for $n \geqslant 0$)

$$\Gamma_{n,m} = \zeta^2 D_m(-in\tau). \tag{10.40}$$

The function $D_m(t)$, starting with $t > 0$, can be analytically continued into the lower half t plane. The reason this is possible is that $H - E_0$ is a positive operator (by definition of E_0) and therefore $\exp\{-i(H - E_0)t\}$ is bounded for $\mathrm{Im}\, t < 0$. This analytic continuation defines $D_m(-it)$. Conversely, if $D_m(-it)$ is known, $D_m(t)$ is determined by analytic continuation.

In summary the spin-spin correlation function of the statistical mechanical theory is equal to the propagator of the lattice quantum theory at discrete values ($n\tau$) of the (imaginary) time variable.

Note that while the original statistical mechanical system was defined on a lattice in the imaginary time direction, the quantum theory derived from the statistical mechanics is defined for a continuous real time variable. However, it is difficult to evaluate $D_m(t)$ for any real values of t or imaginary values between the lattice points in τ. To get to values of t between the lattice points in practice one has to diagonalize V and use expressions like (10.44) below. There are at present no practical methods for diagonalizing V near the critical point (except for the exactly soluble two dimensional Ising model [144]). A study of some eigenstates and eigenvalues of V as an expansion in K is described in [143].

To replace the spatial lattice m by a continuum is more difficult. In the field theoretic derivation of the lattice quantum mechanics, a lattice spacing a was introduced, so one might expect to recover a continuum theory by letting $a \to 0$. Consider, however, the physical (renormalized) mass μ_R of the field theory. If one does not make a mass renormalization, (that is, if μ_0 is held fixed as $a \to 0$) then μ_R will be proportional to the cutoff momentum a^{-1} rather than a constant. This is the conclusion from perturbation theoretic renormalization theory in dimension $d = 4$ (for $3 < d < 4$, μ_R diverges but less rapidly as $a \to 0$). The same conclusion results from looking at the statistical mechanics in the limit of small a, without regard to perturbation theory. As will be shown below, the physical mass is $(\xi a)^{-1}$ where ξ is the correlation length of the statistical mechanical system in units of the lattice spacing. In dimension $d = 4$, ξ most likely stays finite when $a \to 0$ (see below) and hence μ_R behaves as a^{-1}.

To prove that

$$\mu_R = 1/(\xi a) \tag{10.41}$$

one compares the long range behavior of the spin-spin correlation function and the propagator.

Consider the spin-spin correlation function $\Gamma_{n,m}$ in the limit of large n, with $m = 0$:

$$\Gamma_{n,0} \propto \exp(-n/\xi) \tag{10.42}$$

for $n \gg \xi$, apart from a power of n. Now consider the propagator $D_m(-in\tau)$ for $m = 0$ and n large; it is

$$D_0(-in\tau) = \langle\Omega|\phi_0 \exp(-Hn\tau)\,\phi_0|\Omega\rangle \exp(E_0 n\tau). \tag{10.43}$$

Writing this as a sum over all eigenstates $|\nu\rangle$ of H gives,

$$D_0(-in\tau) = \sum_\nu \exp\{-(E_\nu - E_0)n\tau\}\,|\langle\Omega|\phi_0|\nu\rangle|^2. \tag{10.44}$$

For large n the lowest energy dominates. Because of the symmetry for $\phi \to -\phi$, the lowest energy state that contributes is not the ground state but the first excited state. According to conventional wisdom the first excited state is a single particle state at rest. (This can be demonstrated in perturbation theory in K [143] but is not proven generally.) The energy difference $E_\nu - E_0$ for this state is the mass μ_R of the particle. Therefore

$$D_0(-in\tau) \propto \exp(-\mu_R n\tau) \tag{10.45}$$

for $n \to \infty$. (The sum over ν becomes an integral over the momentum of the single particle state and results in a power of n multiplying the exponential; see [147].) Comparing eqs. (10.40), (10.42), and (10.45), one obtains $\mu_R \tau = 1/\xi$. Since we have chosen $\tau = a$, one obtains eq. (10.41).

To see that ξ is likely to be finite for $a \to 0$, consider the statistical mechanics with the parameters given by (10.28), in the limit $a = \tau \to 0$. If ζ is held fixed the parameters b, u_0 and $K = K_1 = K_2$ all go to zero as $a \to 0$, which is not very helpful. For finite a one can choose ζ at will without changing ξ (see below); with the particular choice

$$\zeta = a^{(d-2)/2} \tag{10.46}$$

the parameters b and K have non-zero limits for $a \to 0$:

$$b \to 2d, \qquad K \to 1. \tag{10.47}$$

For any a one has,

$$u_0 = \lambda_0 a^{4-d} \tag{10.48}$$

so $u_0 = \lambda_0$ for $d = 4$ (for $d < 4$, $u_0 \to 0$ as $a \to 0$). The reason that ξ is independent of ζ is that the change from ϕ_m to s_m is only a change of variables; if ζ is changed to ζ' this does not change expectation values of ϕ_m at all and the spin-spin correlation function $\Gamma_{n,m}$ is only changed by the scale factor $(\zeta'/\zeta)^2$. This does not change ξ.

For $d = 4$ and $a \to 0$ the parameters in the statistical mechanics are: $b = 8$, $K = 1$, $u_0 = \lambda_0$. For ξ to be infinite, K must be at its critical value. If $\lambda_0 = 0$, $K = 1$ is the critical value (see section 3), and ξ is infinite. For $\lambda_0 \neq 0$, 1 is not K_c in general and ξ is finite. For example, for large λ_0, $K = 1$ can easily be treated by an expansion in K and is nowhere near the critical point (this is analogous to the expansion proposed by Schiff [148]).

For $d < 4$, u_0 is 0 at $a = 0$, so $K = 1$ is the critical point for any value of λ_0. However, careful examination of how ξ behaves as $a \to 0$ shows that $(\xi a)^{-1}$ is still divergent as $a \to 0$, when d is in the range $3 \leqslant d < 4$.

This ends the digression.

The basic aim of this section is to construct a field theory from the statistical mechanics. In this case one has no a priori rules for how b, u_0, and K should depend on the lattice spacing a. No

such constant appears in the statistical mechanics because one naturally expresses lengths in units of the lattice spacing. In field theory one has to introduce the lattice spacing a as a nontrivial parameter because for a field theorist the natural unit of length is the reciprocal of μ_R. Even if one does not set $\mu_R = 1$, it is necessary for μ_R to have a finite limit when $a \to 0$.

In preparation for constructing the field theory, it is logical to discuss the statistical mechanics in units with $\mu_R = 1$. These units can be used for any choice of the statistical mechanical parameters b, u_0, and K not at the critical point. The lattice spacing in these units is $a = 1/\xi(b, u_0, K)$. It is a somewhat peculiar change of units since it depends on the parameters b, u_0 and K. Working in these units, μ_R remains fixed no matter how one varies the parameters. In particular, the limit $a \to 0$ holding μ_R fixed is obtained by letting $K \to K_c(b, u_0)$ where $K_c(b, u_0)$ is the critical value of K.

Thus it is rather trivial to obtain a finite mass μ_R in the limit $a \to 0$. What is less trivial is to obtain definite limits for the field theoretic vacuum expectation values as $a \to 0$. The trouble is that one has to know how the multi-spin correlation functions behave as $K \to K_c$. In particular, one is interested in vacuum expectation values for distances which are fixed in units of μ_R^{-1}. A fixed distance in these units is a distance proportional to ξ in units of the lattice spacing. Thus, if the long range correlation functions have too complicated a behavior as $K \to K_c$, there will be no definite limit for the vacuum expectation values as $K \to K_c$. In section 12 it will be shown that the continuum limit does exist for all vacuum expectation values if the critical behavior is determined by a fixed point of the renormalization group.

We conclude with two brief comments. For $d < 4$ the conventionally defined $\lambda_0 \phi^4$ theory is easy to renormalize in perturbation theory; only a mild mass renormalization is required. This is related to the fact that $u_0 \to 0$ as $a \to 0$ as discussed earlier. However, one can construct a more general field theory for $d < 4$ starting from the statistical mechanics and holding u_0 fixed (instead of λ_0) as $a \to 0$. The existence of the limit in this case is established (assuming the existence of a fixed point) in section 12. It is this theory that has tensor operators with the anomalous dimensions calculated in section 9.

The second comment is this. One normally takes for granted that the Hilbert space of a quantum theory has a positive metric (although there has been some flirting with quantum field theories with indefinite metric by Heisenberg and others). The use of the transfer matrix formalism insures that the statistical mechanics defines a lattice quantum theory with a positive metric. This is a nontrivial result, because one can define statistical mechanical models which *cannot* be equivalent to quantum theories with a positive metric. A simple example is a Gaussian model with interaction,

$$\mathcal{H}_0 = \int_q [r_0 + q^2]^2 \sigma_q \sigma_{-q}. \tag{10.49}$$

The propagator of this theory is $1/(r_0 + q^2)^2$ which has a double pole at $q^2 = -r_0$. A double pole cannot occur for a quantum theory with positive metric. Therefore, this theory cannot be written as a lattice theory having a transfer matrix. The main restriction of the transfer matrix approach is that there only be nearest neighbor interactions between rows, so that eq. (10.7) is valid. Hence, lattice approximations to (10.49) necessarily involve second nearest neighbor interactions at least.

11. Exact renormalization group equations in differential form

The purpose of this section is to present an exact formulation of the renormalization group in differential form. These equations are very complicated so they will not be discussed in great

detail. However, the underlying concepts will be stressed. (The equations derived here were reported informally by KGW to the Irvine conference (1970); they have not been published previously.)

The formal discussion of consequences of the renormalization group works best if one has a differential form of the renormalization group transformation. Also, a differential form is useful for the investigation of properties of the ϵ expansion to all orders (for questions like "does it exist"; this question has not been answered yet). A longer range possibility is that one will be able to develop approximate forms of the transformation which can be integrated numerically; if so, one might be able to solve problems which cannot be solved any other way.

In previous sections a discrete renormalization group transformation was defined by integrating out the spin components σ_q with $|q| > \frac{1}{2}$. Then one made a change of scale in the remaining σ_q ($|q| < \frac{1}{2}$) replacing them by $\zeta\sigma'_{2q}$ with ζ being a constant.

The obvious way to construct an infinitesimal transformation is to integrate out only those σ_q with $1 - \delta t < |q| < 1$ for δt small. This however turns out to be difficult. The problem is that normally to order δt a Feynman diagram will have at most one internal momentum in this range; however, when two internal momenta are restricted to be $+q$ and $-q$ by particular choices of the external momenta (e.g., when $p = 0$ in fig. 11.1), then two internal momenta can be simultaneously in the range $1 - \delta t < |q| < 1$. This fact creates problems. See supplemental ref. [5] for further discussion.

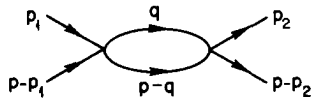

Fig. 11.1. A Feynman diagram which is difficult to handle in an infinitesimal formulation of the discrete renormalization group transformation.

More philosophically speaking it is desirable to avoid sharp boundaries in momentum space between integrated and unintegrated spin components σ_q. The trouble is that sharp boundaries in momentum space result in non-local interactions in position space which one would prefer to avoid (see section 1).

To avoid sharp boundaries a trick will be used. A smooth interpolation procedure between integrated and unintegrated variables will be introduced, called "incomplete integration". This will permit replacing the sharp boundary in q space by a finite interval in q for which σ_q is "incompletely" integrated. (See fig. 11.2.) There are many ways of realizing the idea of incomplete integration. The particular choice presented below happens to have many useful features but

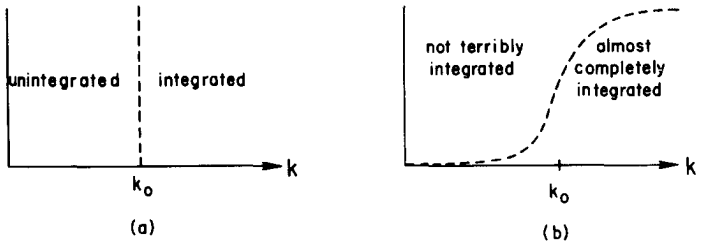

Fig. 11.2. (a) A sharp boundary between integrated and unintegrated momenta. (b) A smooth boundary between almost completely integrated and not terribly integrated momenta.

is certainly not unique. The idea will be introduced by considering an ordinary integral in place of a functional integral.

Recall a few facts about differential equations. In particular, consider the differential equation,

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} + x\right)\psi. \tag{11.1}$$

Solutions to (11.1) can be written in terms of a Green's function,

$$\psi(x', t) = \int \mathrm{d}x\, G(x', t; x, 0)\psi_0(x). \tag{11.2}$$

Because (11.1) is quadratic in x and $\partial/\partial x$, the Green's function is known explicitly: it is,

$$G(x', t; x, 0) = \frac{1}{\sqrt{2\pi(1 - \exp(-2t))}} \exp\left[-\frac{(x' - x\mathrm{e}^{-t})^2}{2(1 - \mathrm{e}^{-2t})}\right]. \tag{11.3}$$

Eq. (11.2) provides a particularly simple realization of the notion of incomplete integration. As $t \to 0$, the Green's function becomes $\delta(x - x')$ and $\psi(x', t)$ becomes the initial arbitrary function $\psi_0(x')$,

$$\psi(x', 0) = \psi_0(x'). \tag{11.4}$$

As t becomes very large $\psi(x', t)$ becomes proportional to the integral $\int \psi_0(x)\,\mathrm{d}x$ apart from the known Gaussian function $\exp(-\frac{1}{2}x'^2)$,

$$\psi(x', \infty) = \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}x'^2) \int_{-\infty}^{\infty} \psi_0(x)\,\mathrm{d}x. \tag{11.5}$$

Therefore, considered as a function of t, $\psi(x', t)$ gives an interpolation between a completely unintegrated function $\psi_0(x')$ and its completely integrated form. Thus t gives a measure of how far the variable x has been integrated.

Consider now the functional integral. Since each momentum component of the fields must be integrated over, a quantity $\alpha_q(t)$ must be introduced to measure how completely each component σ_q has been integrated. A particular choice for $\alpha_q(t)$ will be introduced shortly. In general, one wants $\alpha_q(t)$ to have the shape displayed in fig. 11.3. Let σ_q'' denote the spin variable for the Hamiltonian $\mathcal{H}_t$ before the momentum has been rescaled from the original momentum scale. The

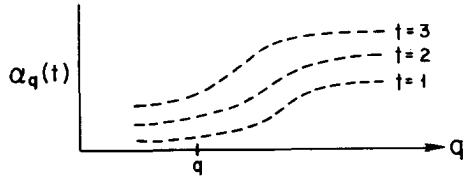

Fig. 11.3. A choice of the function $\alpha_q(t)$ which measures (as a function of t) how completely each component σ_q has been integrated.

spin variable after the scale change will be denoted σ'_q. The functional analogues of (11.1) and (11.2) become,

$$\frac{\delta}{\delta\alpha_q}\exp\{\mathcal{H}_t[\sigma'']\} = \frac{\delta}{\delta\sigma''_q}\left(\frac{\delta}{\delta\sigma''_{-q}} + \sigma''_q\right)\exp\{\mathcal{H}_t[\sigma'']\}, \tag{11.6}$$

$$\exp\{\mathcal{H}_t[\sigma'']\} = c\int_\sigma \exp\left\{-\tfrac{1}{2}\int_q \frac{(\sigma''_q - \sigma_q\exp\{-\alpha_q(t)\})(\sigma''_{-q} - \sigma_{-q}\exp\{-\alpha_q(t)\})}{1-\exp\{-2\alpha_q(t)\}}\right\}\exp\{\mathcal{H}_0[\sigma]\}. \tag{11.7}$$

The constant c is unimportant since one only computes ratios of expectation values involving $\exp(\mathcal{H}_t)$. σ''_{-q} and σ_{-q} appear in the right-hand side of (11.7) in order that $\mathcal{H}_t[\sigma]$ conserves momentum. A functional differential equation for $\mathcal{H}_t[\sigma'']$ follows straightforwardly from (11.6)

$$\frac{\partial\mathcal{H}_t}{\partial t} = \int_q \frac{\partial\alpha_q(t)}{\partial t}\left\{\frac{\delta\mathcal{H}_t}{\delta\sigma''_q}\frac{\delta\mathcal{H}_t}{\delta\sigma''_{-q}} + \frac{\delta^2\mathcal{H}_t}{\delta\sigma''_q\,\delta\sigma''_{-q}} + \sigma''_q\frac{\delta\mathcal{H}_t}{\delta\sigma''_q} + \text{const.}\right\}. \tag{11.8}$$

The "const." is a σ'' independent term which can be ignored.

Now a useful $\alpha_q(t)$ must be chosen. This function should have the property that it separates small q from large q. A good choice turns out to be,

$$\alpha_q(t) = q^2(e^{2t}-1) + \rho(t), \qquad \rho(0) = 0. \tag{11.9}$$

The purpose of the function $\rho(t)$ is to allow a normalization condition to be imposed on the kinetic term in $\mathcal{H}_t$, say

$$\mathcal{H}_t = -\tfrac{1}{2}\int_{q'} q'^2\sigma'_{q'}\sigma'_{-q'} + \ldots \tag{11.10}$$

(in terms of $\sigma'_{q'}$ which will be defined shortly). Clearly $\rho(t)$ serves the same purpose that ζ did in the discrete scale changes of earlier chapters (see, e.g. eq. (3.33)). Some such normalization condition turns out to be essential if critical behavior is to be described by a fixed point. Imposing this normalization condition on $\mathcal{H}_t$ for all t results in an equation for $\rho(t)$.

The final step in constructing the renormalization group transformation is the change of scale, i.e. $\mathcal{H}_t$ should be written in terms of a momentum variable q' whose range is effectively $0 < |q'| < 1$. From (11.7) and (11.9) it is clear that the effective cutoff on q is e^{-t}, so the rescaled momentum q' should be $q' = qe^t$. Also, the rescaled spin variable $\sigma'_{q'}$ must be introduced. It will be defined to be

$$\sigma'_{q'} = \exp(-dt/2)\,\sigma''_q. \tag{11.11}$$

The scale factor is determined by the requirement that the differential renormalization group transformation we will derive must be independent of t. (It is easily seen from (11.8) that one cannot make arbitrary changes of normalization of σ''_q without fouling up the equation.)

The functional differential equation must now be rewritten with $\sigma'_{q'}$ replacing σ''_q. The time derivative $\partial\mathcal{H}_t/\partial t$ in (11.8) is for a fixed function σ''_q. This must be replaced by a derivative for fixed $\sigma'_{q'}$. The relation between the two time derivatives is

$$\left.\frac{\partial\mathcal{H}_t}{\partial t}\right|_{\sigma''} = \left.\frac{\partial\mathcal{H}_t}{\partial t}\right|_{\sigma'} + \int_{q'}\left.\frac{d\sigma'_{q'}}{dt}\right|_{\sigma''}\frac{\delta\mathcal{H}_t}{\delta\sigma'_{q'}}. \tag{11.12}$$

The derivative $\mathrm{d}\sigma'_{q'}/\mathrm{d}t|_{\sigma''}$ follows from (11.11),

$$\left.\frac{\mathrm{d}\sigma'_{q'}}{\mathrm{d}t}\right|_{\sigma''} = \left(-\frac{d}{2} - \boldsymbol{q}' \cdot \nabla_{q'}\right)\sigma'_{q'} \tag{11.13}$$

so (11.8) becomes,

$$\left.\frac{\partial \mathcal{H}_t}{\partial t}\right|_{\sigma'} = \int_{q'}\left[\left(\frac{d}{2} + \boldsymbol{q}' \cdot \nabla_{q'}\right)\sigma'_{q'}\right]\frac{\delta \mathcal{H}_t}{\delta \sigma'_{q'}} + \int_q \frac{\partial \alpha_q}{\partial t}\left[\frac{\delta \mathcal{H}_t}{\delta \sigma''_q}\,\frac{\delta \mathcal{H}_t}{\delta \sigma''_{-q}} + \frac{\delta^2 \mathcal{H}_t}{\delta \sigma''_q\,\delta \sigma''_{-q}} + \sigma''_q\,\frac{\delta \mathcal{H}_t}{\delta \sigma''_q}\right]. \tag{11.14}$$

But, from (11.9)

$$\partial \alpha_q/\partial t = 2q^2 \mathrm{e}^{2t} + \mathrm{d}\rho/\mathrm{d}t = 2q'^2 + \mathrm{d}\rho/\mathrm{d}t. \tag{11.15}$$

Finally, (11.11) and properties of functional derivatives imply,

$$\delta \mathcal{H}_t/\delta \sigma''_q = \exp(dt/2)\,\delta \mathcal{H}_t/\delta \sigma'_{q'}. \tag{11.16}$$

Substituting (11.15) and (11.16) into (11.14) gives the desired renormalization group differential equation,

$$\frac{\partial \mathcal{H}_t}{\partial t} = \int_{q'}\left(\frac{d}{2}\sigma'_{q'} + \boldsymbol{q}' \cdot \nabla_{q'}\sigma'_{q'}\right)\frac{\delta \mathcal{H}_t}{\delta \sigma'_{q'}} + \int_{q'}\left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q'^2\right)\left(\frac{\delta \mathcal{H}_t}{\delta \sigma'_{q'}}\,\frac{\delta \mathcal{H}_t}{\delta \sigma'_{-q'}} + \frac{\delta^2 \mathcal{H}_t}{\delta \sigma'_{q'}\,\delta \sigma'_{-q'}} + \sigma'_{q'}\,\frac{\delta \mathcal{H}_t}{\delta \sigma'_{q'}}\right). \tag{11.17}$$

This is a functional differential equation. It can be reduced to ordinary partial differential equations by expanding $\mathcal{H}_t$ in powers of σ: write (the primes on σ and $\boldsymbol{q}$ are omitted),

$$\mathcal{H}_t[\sigma] = -\tfrac{1}{2}\int_q u_2(\boldsymbol{q}, t)\sigma_{\boldsymbol{q}}\,\sigma_{-\boldsymbol{q}} - \frac{1}{4!}\int_q\int_{q_1}\int_{q_2} u_4(\boldsymbol{q}, \boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, t)\sigma_{\boldsymbol{q}}\sigma_{\boldsymbol{q}_1}\sigma_{\boldsymbol{q}_2}\sigma_{\boldsymbol{q}_3} - \ldots \tag{11.18}$$

(where $\boldsymbol{q}_3 = -\boldsymbol{q} - \boldsymbol{q}_1 - \boldsymbol{q}_2$ to conserve momentum). A constant independent of σ has been omitted. The equations for u_2, u_4, and u_6 are

$$\frac{\partial u_2}{\partial t}(\boldsymbol{q}, t) = \left\{-\boldsymbol{q} \cdot \nabla + 2\left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q^2\right)[1 - u_2(\boldsymbol{q}, t)]\right\}u_2(\boldsymbol{q}, t) + \int_{q_1}\left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q_1^2\right)u_4(\boldsymbol{q}, -\boldsymbol{q}, \boldsymbol{q}_1, -\boldsymbol{q}_1, t)$$

$$\frac{\partial u_4}{\partial t}(\boldsymbol{q} \ldots \boldsymbol{q}_3) = \left\{-d - \boldsymbol{q} \cdot \nabla_{\boldsymbol{q}} + \left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q^2\right)[1 - 2u_2(\boldsymbol{q}, t)] - \ldots - \boldsymbol{q}_3 \cdot \nabla_{\boldsymbol{q}_3} + \left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q_3^2\right)\right.$$
$$\left.\times\,[1 - 2u_2(\boldsymbol{q}_3, t)]\right\}u_4(\boldsymbol{q}, \ldots, \boldsymbol{q}_3, t) + \int_{q_4}\left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q_4^2\right)u_6(\boldsymbol{q}, \boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4, -\boldsymbol{q}_4, t)$$

$$\frac{\partial u_6}{\partial t}(\boldsymbol{q} \ldots \boldsymbol{q}_5, t) = \left\{-2d - \boldsymbol{q} \cdot \nabla_{\boldsymbol{q}} + \left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q^2\right)[1 - 2u_2(\boldsymbol{q}, t)] - \ldots - \boldsymbol{q}_5 \cdot \nabla_{\boldsymbol{q}_5} + \left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q_5^2\right)\right.$$
$$\left.\times\,[1 - 2u_2(\boldsymbol{q}_5, t)]\right\}u_6(\boldsymbol{q}, \ldots, \boldsymbol{q}_5, t) - 2\left[\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2(\boldsymbol{q} + \boldsymbol{q}_1 + \boldsymbol{q}_2)^2\right]\cdot$$
$$\times\,u_4(\boldsymbol{q}, \boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3 + \boldsymbol{q}_4 + \boldsymbol{q}_5, t)\,u_4(\boldsymbol{q}_3, \boldsymbol{q}_4, \boldsymbol{q}_5, \boldsymbol{q} + \boldsymbol{q}_1 + \boldsymbol{q}_2, t) - (9\text{ permutations})$$
$$+ \int_{q_6}\left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q_6^2\right)u_8(\boldsymbol{q}, \ldots, \boldsymbol{q}_5, \boldsymbol{q}_6, -\boldsymbol{q}_6, t). \tag{11.19}$$

The "9 permutations" in the equation for u_6 means the 9 inequivalent permutations of $(q_1, \ldots, q_5)$ in the $u_4 \times u_4$ term.

So far, no equation has been given for $\mathrm{d}\rho/\mathrm{d}t$. The normalization condition (11.10) means that if $u_2(q, t)$ is expanded in powers of q^2:

$$u_2(q, t) = u_{2A}(t) + u_{2B}(t)q^2 + \ldots \tag{11.20}$$

then $u_{2B}(t) = 1$ independent of t.

This means $\mathrm{d}u_{2B}(t)/\mathrm{d}t$ is 0, which results in an equation for $\mathrm{d}\rho/\mathrm{d}t$. Write u_4 as an expansion in powers of q:

$$u_4(q, -q, q_1, -q_1, t) = u_{4A}(q_1, t) + q \cdot q_1 u_{4B}(q_1, t) + q^2 u_{4C}(q_1, t) + (q \cdot q_1)^2 u_{4D}(q_1, t) + \ldots . \tag{11.21}$$

(This expansion assumes rotational symmetry.) Then the equation determining $\mathrm{d}\rho/\mathrm{d}t$ is,

$$0 = 2u_{2B}\frac{\mathrm{d}\rho}{\mathrm{d}t}(t)\,[1 - 2u_{2A}(t)] - 2u_{2B} + 4u_{2A}[1 - u_{2A}] + \int_{q_1}\left[\frac{\mathrm{d}\rho}{\mathrm{d}t}(t) + 2q_1^2\right] \times [u_{4C}(q_1, t) + d^{-1}q_1^2 u_{4D}(q_1, t)]. \tag{11.22}$$

However the reader is warned that the simple normalization condition $u_{2B}(t) = 1$ for all t cannot always be realized; one must be prepared to allow $u_{2B}(t)$ to be another constant or perhaps t dependent. See the Appendix for an example of the solution of the renormalization group equations and a determination of $\mathrm{d}\rho/\mathrm{d}t$.

We know how to obtain spin-spin correlation functions given $\mathcal{H}_0$, but it is desirable to be able to obtain them when given $\mathcal{H}_t$. To do this first consider the generating functional originally given in terms of $\mathcal{H}_0$ and then rewrite it in terms of $\mathcal{H}_t$. The analogous problem for one variable, x, will be sketched to begin. The $\psi_0(x)$ of eq. (11.2) is analogous to $\exp(\mathcal{H}_0[\sigma])$ so construct the generating functional:

$$Z(j) = \int_{-\infty}^{\infty} \psi_0(x') \exp(jx')\,\mathrm{d}x'. \tag{11.23}$$

A straightforward calculation shows that this is equivalent to

$$Z(j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{jx'\mathrm{e}^t - \frac{j^2(\mathrm{e}^{2t} - 1)}{2}\right\} G(x', t; x, 0)\,\psi_0(x)\,\mathrm{d}x'\,\mathrm{d}x \tag{11.24}$$

for any t. (To prove this one performs the x' integration using the explicit form (11.3) for the Green's function.) Reversing the order of integration gives

$$Z(j) = \int_{-\infty}^{\infty} \psi(x', t) \exp\left\{jx'\mathrm{e}^t - \frac{j^2(\mathrm{e}^{2t} - 1)}{2}\right\}\mathrm{d}x'. \tag{11.25}$$

Now consider the generating functional of actual interest

$$Z(j) = \int_{\sigma} \exp\left\{\int_q j(q)\sigma_{-q} + \mathcal{H}_0\right\}. \tag{11.26}$$

In terms of $\mathcal{H}_t$ and σ'', one has in analogy to (11.25),

$$Z(j) = \int_{\sigma''} \exp\left\{ \sum_q j(q)\sigma''_{-q} \exp\{\alpha_q(t)\} - \tfrac{1}{2} \sum_q j(q)j(-q)[\exp\{2\alpha_q(t)\} - 1] + \mathcal{H}_t \right\} \tag{11.27}$$

and converting to σ' gives

$$Z(j) = \int_{\sigma'} \exp\left\{ \sum_{q'} e^{-dt/2} e^{\rho(t)} j(q'e^{-t})\sigma_{q'} \exp[(q')^2 - (q')^2 e^{-2t}] \right.$$
$$\left. - \tfrac{1}{2} \sum_{q'} j(q'e^{-t})j(-q'e^{-t})e^{-dt}[\exp(2q'^2 - 2q'^2 e^{-2t} + 2\rho(t)) - 1] + \mathcal{H}_t[\sigma'] \right\}. \tag{11.28}$$

This is a rather complicated formula. According to the general principles of the renormalization group approach, the only change in calculating correlation functions from $\mathcal{H}_t$ as opposed to $\mathcal{H}_0$ should be a scale change; namely, σ_{-q} should be replaced by $\zeta_t \sigma'_{-q'}$ with $q' = qe^t$. In other words, one should write

$$\sum_q j(q)\sigma_{-q} = e^{-dt}\zeta_t \sum_{q'} j(q'e^{-t})\sigma'_{-q'} = \sum_{q'} j(q', t)\sigma'_{-q'} \tag{11.29}$$

with

$$j(q', t) = e^{-dt}\zeta_t j(q'e^{-t}). \tag{11.30}$$

Suppose one defines

$$\zeta_t = e^{dt/2} e^{\rho(t)}, \tag{11.31}$$

then (11.28) becomes (dropping primes for convenience),

$$Z(j) = \int_{\sigma} \exp\left\{ \sum_q j(q, t)\sigma_q \exp[q^2 - q^2 e^{-2t}] - \tfrac{1}{2} \sum_q j(q, t)j(-q, t)[\exp(2q^2 - 2q^2 e^{-2t}) - e^{-2\rho(t)}] \right.$$
$$\left. + \mathcal{H}_t[\sigma] \right\}. \tag{11.32}$$

This formula for computing $Z(j)$ is more complicated than expected from general principles. However, it will not affect the general renormalization group analysis: For large t and q of order 1, the e^{-2t} and $e^{-2\rho(t)}$ terms can be neglected (neglecting $e^{-2\rho(t)}$ assumes that $\rho(t) \to \infty$ as $t \to \infty$; this is true of examples that have been studied to date). Then $Z(j)$ depends on t only through $j(q, t)$ and the t dependence of $\mathcal{H}_t$. This is the crucial fact needed for a general discussion of consequences of the renormalization group.

It is remarkable that one can write an exact yet completely explicit form for $\partial\mathcal{H}_t/\partial t$ (eq. (11.17)). Previously all formulations of the renormalization group have involved perturbation theory in an essential way (as in the Gell-Mann–Low formulation or the graphical formulation of sections 4 and 5) or else involve incalculable functions (the Kadanoff approach). Eq. (11.17) can be nasty to work with, but in the long run it or similar equations will probably be the basis for most work on the renormalization group.

For future reference the form of $Z(j)$ for large t is,

$$Z(j) = \int_{\sigma} \exp\left\{ \sum_q j(q, t)\sigma_{-q} \exp(q^2) - \tfrac{1}{2} \sum_q j(q, t)j(-q, t) \exp(2q^2) + \mathcal{H}_t[\sigma] \right\}. \tag{11.33}$$

This is the formula that is used for $\mathcal{H}_t[\sigma]$ near a fixed point $\mathcal{H}^*$ or on a renormalized trajectory leaving $\mathcal{H}^*$.

Note that actual correlation functions always involve a denominator Z^{-1} so the generating functional of practical interest is $Z(j)/Z(0)$. This is fortunate because the constant term in $\mathcal{H}_t$ was neglected in deriving the differential equation (11.17).

There is another useful formula for the generating functional $Z(j)$ (more precisely, $Z(j)/Z(0)$). The formula is the following. Define a particular function $\sigma'_t(q)$ to be

$$\sigma'_t(q) = j(q\mathrm{e}^{-t}) \exp\{\beta_q(t) - dt/2\} \tag{11.34}$$

where

$$\beta_q(t) = q^2(1 - \mathrm{e}^{-2t}) + \rho(t) = \alpha_{q\ \exp(-t)}(t). \tag{11.35}$$

Then,

$$Z(j) = \lim_{t\to\infty} \exp\left\{\mathcal{H}_t[\sigma'_t] + \tfrac{1}{2}\int_q \sigma'_t(q)\sigma'_t(-q)[1 + \exp\{-2\beta_q(t)\}]\right\}. \tag{11.36}$$

This formula is demonstrated as follows. First write (11.7) in terms of $\sigma'_t(q)$, then let $t\to\infty$ and use the definition (11.26) to identify $Z(j)$. To begin relate the unscaled variable σ''_t to the scaled variable σ'_t,

$$\sigma''_t(q) = \exp(dt/2)\ \sigma'_t(q\mathrm{e}^t) = j(q)\exp\{\alpha_q(t)\}. \tag{11.37}$$

Substitution of $\sigma''_t(q)$ for σ''_q in (11.7) gives,

$$\exp\{\mathcal{H}_t[\sigma'_t]\} = c\exp\left\{-\tfrac{1}{2}\int_q \frac{j(q)j(-q)\exp\{2\alpha_q(t)\}}{1-\exp\{-2\alpha_q(t)\}}\right\}$$
$$\times \int_\sigma \exp\left\{\int_q \frac{j(q)\sigma_{-q}}{1-\exp\{-2\alpha_q(t)\}} - \tfrac{1}{2}\int_q \frac{\sigma_q\sigma_{-q}\exp\{-2\alpha_q(t)\}}{1-\exp\{-2\alpha_q(t)\}} + \mathcal{H}_0[\sigma]\right\}. \tag{11.38}$$

In the limit $t\to\infty$, (11.38) can be rewritten,

$$\lim_{t\to\infty} \exp\left\{\mathcal{H}_t[\sigma'_t] + \tfrac{1}{2}\int_q \sigma'_t(q)\sigma'_t(-q)[1 + \exp\{-2\beta_q(t)\}]\right\} = c\int_\sigma \exp\left\{\int_q j(q)\sigma_{-q} + \mathcal{H}_0[\sigma]\right\}. \tag{11.39}$$

From its definition (11.26), $Z(j)$ can be recognized as (11.39). The advantage of (11.36) is that it avoids the explicit calculation of functional integrals; it is only necessary to solve the functional differential equation for $\mathcal{H}_t$.

12.

12.1. *Topology of the renormalization group transformation (fixed points, trajectories, and subspaces)*

Underlying the detailed calculations and analyses of the renormalization group approach described in previous sections are some simple topological ideas. Fixed points, for example, have already been emphasized. In this section topological ideas will be considered at length. They will

be discussed in simple but abstract form without reference to specific forms of the renormalization group transformation. A few general assumptions will be made concerning the effective interactions generated by the renormalization group. These assumptions are true of solutions (exact or approximate) obtained to date, but have not been proven in general. Simple forms of the topology will be considered to illustrate the ideas. In actual examples the topology gets more complicated (e.g. many different fixed points can occur) but one can handle complications once the ideas have been mastered.

Fixed points were discussed in the original paper of Gell-Mann and Low [149] (see also [150]). The need for a full topological discussion arose only when the renormalization group transformation was defined to transform many coupling constants rather than just one. A complete topological analysis was first given for a simplified fixed source model [151]; much of the inspiration for this section comes from [151]. Wegner [152] has given an extensive discussion of solutions of the renormalization group near a fixed point, which overlaps and considerably extends the last part of this section.

Spaces and subspaces

A renormalization group transformation is a transformation U on a space of cutoff interactions S. In the previous lecture an example of an infinitesimal transformation U has been given. The renormalization group equation is

$$\partial \mathcal{H}_t / \partial t = U[\mathcal{H}_t]. \tag{12.1}$$

The interactions $\mathcal{H}_t$ come from the space S.

The interactions in S are given in dimensionless form with momenta given in units of the cutoff. One is always free to take any interaction $\mathcal{H} \in S$ and change the cutoff by making a change of units. The cutoff can consist of an upper bound on all momenta q imposed either directly or through the introduction of a lattice (cf. sections 3 and 10). The cutoff is necessary because the renormalization group transformation U is defined to integrate out the momenta just below the cutoff and that is a meaningless operation if the cutoff is infinite. In the case of the exact renormalization group equations of section 11, a strict upper bound on q is not necessary, but it is difficult to integrate the equations if the non-Gaussian terms in $\mathcal{H}_0$ are not cutoff for momenta $q \gg 1$. In the renormalization group approach an uncutoff field theory has to be obtained as the limit of cutoff theories. (This is independent of whether one is calculating in perturbation theory or not.)

The coordinates in S are the free parameters in an arbitrary interaction $\mathcal{H} \in S$. In the simple examples considered in previous lectures, r and u were coordinates in S. To discuss the exact renormalization group equations, an arbitrary interaction in $\mathcal{H}$ is defined in terms of an infinite set of arbitrary functions $u_2(q)$, $u_4(q, q_1, q_2, q_3)$, etc. This means S is an infinite dimensional space.

A solution $\mathcal{H}_t$ of the renormalization group, as a function of t, provides a set of interactions all of which describe the same physical system. However one must make a t-dependent change of scale when relating $\mathcal{H}_t$ to the physics. One can first of all make a scale change when going from $\mathcal{H}_0$ to the physics so that the physics has a cutoff momentum $\Lambda_0 \neq 1$. This is a scale change with no restrictions: Λ_0 can be chosen at will. Given Λ_0, the scale change which relates $\mathcal{H}_t$ to the physics is fixed, namely the cutoff momentum of $\mathcal{H}_t$ in physical units must be $e^{-t}\Lambda_0$. (See, e.g., section 7.)

Let the physical system have a correlation length ξ_p in physical units (cm or MeV^{-1}, say). This must be independent of which interaction $\mathcal{H}_t$ (i.e. which value of t) is used to describe the physics.

Let ξ_t be the correlation length of $\mathcal{H}_t$ in dimensionless units. Then the scale changes defined above imply that

$$\xi_t = e^{-t}\xi_p\Lambda_0. \tag{12.2}$$

This in particular means that ξ_t satisfies

$$\xi_t = e^{-t}\xi_0 \tag{12.3}$$

for a solution $\mathcal{H}_t$ of eq. (12.1).

The interaction $\mathcal{H}_t$ rescaled to physical units has an artificially small cutoff $e^{-t}\Lambda_0$ compared to the true cutoff Λ_0 of the physical system. This is by construction; the higher momenta in $\mathcal{H}_0$ have been removed by integration. The artificially low cutoff means that the interactions in $\mathcal{H}_t$ are artificially of long range. The interactions in the rescaled $\mathcal{H}_t$ will have a range of order e^t/Λ_0, at least. The initial interaction $\mathcal{H}_0$, rescaled, has interactions only of range $1/\Lambda_0$ unless one deliberately adds longer range interactions.

In practice a "short range" term in $\mathcal{H}_0$, before rescaling, is for example $\int_x\int_y \exp\{-(x-y)^2\} \times s(x)s(y)$ while examples of "long range" interactions would be $\int_x\int_y \exp\{-0.0001(x-y)^2\}\, s(x)s(y)$ (which has a long but finite range) or $\int_x\int_y |x-y|^{-1}s(x)s(y)$ (which has infinite range).

The space S will be restricted to interactions of short range in dimensionless units. There are several reasons for this. One is a practical one: the approximation schemes of previous lectures considered so far assume that $\mathcal{H}_t$ has only simple short range interactions. A more fundamental reason is that the qualitative ideas about critical behavior (like universality) are known to be false if long range interactions are permitted. (See, e.g., [153].) It will also be crucial to the renormalization theory described later that S have only short range terms.

The space S must be large enough so that for any interaction $\mathcal{H}_0 \in S$, all the effective interactions $\mathcal{H}_t$ generated from $\mathcal{H}_0$ lie in S. Thus one must verify, for example, that the interactions $\mathcal{H}_t$ are all of short range (before rescaling). There is unfortunately no certainty that this is the case in general. It is true of examples solved so far. (Note however that an incorrect choice of $\rho(t)$ in the equations of section 11 can lead to long range interactions: see the Appendix.) In approximate formulations of the renormalization group (e.g., the approximate recursion formula of section 6) one can replace S by a subspace of S but one must make sure that the approximate transformation transforms the subspace into itself.

There are four types of subspaces or "surfaces" in S which should be defined. First the elements of S can be separated according to their correlation lengths. Thus, one defines surfaces consisting of all interactions $\mathcal{H} \in S$ with given dimensionless correlation length ξ. Of special interest will be the "critical surface" with $\xi = \infty$ denoted S_c. Fig. 12.1 shows a possible set of surfaces of fixed ξ in a space parametrized by r and u only.

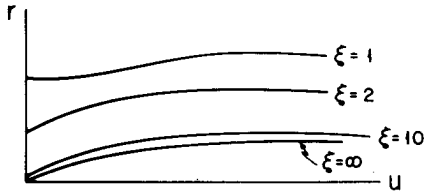

Fig. 12.1. Lines characterized by fixed values of the correlation length ξ. The line labelled $\xi = \infty$ is the "critical surface" S_c.

Secondly, there is a surface of simple interactions to be called the "canonical surface". This surface consists of the set of initial interactions $\mathcal{H}_0$ for a given system or model. For example, the canonical surface might consist of the interactions (4.2) with r and u as the only free parameters.

Thirdly, there are surfaces of definite symmetry. For example, in order to allow for interactions on a lattice, S includes non-rotationally invariant interactions. (Interactions on a three-dimensional cubic lattice have cubic symmetry but not full rotational symmetry.) Then there will be a subspace S_R of interactions which are rotationally invariant.

Finally there are the subspace $S(t)$ generated by applying U to the space S itself. This means the following. Suppose one starts with a space S as illustrated in fig. 12.2 and an infinitesimal

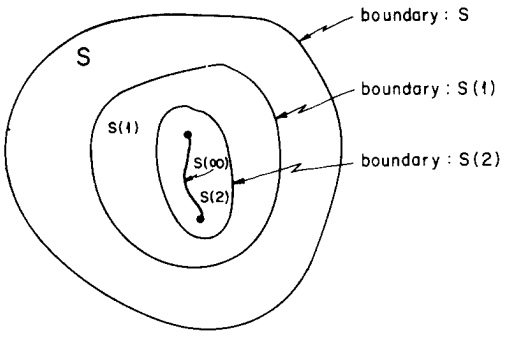

Fig. 12.2. Nested subspaces $S(t)$. In this example the subspace $S(\infty)$ is one dimensional.

transformation U defined on S. The space $S(\delta t)$ with δt infinitesimal is defined as the set of points $\mathcal{H}'$ such that,

$$\mathcal{H}' = \mathcal{H} + \delta t \cdot U[\mathcal{H}] \tag{12.4}$$

where $\mathcal{H}$ is some point in S. One then obtains the space $S(2\delta t)$ by applying U to the space $S(\delta t)$: $S(2\delta t)$ consists of all interactions $\mathcal{H}'$ of the form (12.2) where $\mathcal{H}$ is some point in $S(\delta t)$. By induction one can define $S(t)$ for any t. In other words, if one solves the renormalization group equations to obtain $\mathcal{H}_t$ from $\mathcal{H}_0$, then $S(t)$ is the space of all interactions $\mathcal{H}_t$ generated from all possible initial interactions $\mathcal{H}_0$ contained in S.

By assumption the effective interactions $\mathcal{H}_t$ lie in S so $S(t)$ is a subspace of S.

The subspaces $S(t)$ are nested: $S(t + t') \subset S(t)$ for any positive t and t'. The reason is simple: $S(t + t')$ is obtained by integrating the transformation U through a time $t + t'$, starting from all points in S. This is equivalent to integrating U through a time t starting from all points in $S(t')$. Since $S(t')$ is contained in S, one gets a larger space by starting from all points in S instead of $S(t')$: hence, $S(t + t') \subset S(t)$.

The spaces $S(t)$ for finite t have the same dimensionality as S itself. However, the space $S(\infty)$ (obtained as the limit of $S(t)$ for $t \to \infty$) can be smaller. A basic idea of the renormalization group approach is that $S(\infty)$ is a surface of many less dimensions than S. For illustration the surface $S(\infty)$ in fig. 12.2 is a single curve C. The reason that $S(\infty)$ is expected to be much smaller than S is the following: $\mathcal{H}_t$ (for large t) describes a physical system with forces of a much shorter range than the range of interactions in $\mathcal{H}_t$ itself (as shown earlier). It is unlikely that an arbitrary interaction $\mathcal{H}$ in S has this property; thus $S(t)$ for large t should be only a small subset of S.

A differential equation like (12.1) can be integrated backwards in t as well as forwards.

Suppose $\mathcal{H}$ is in S but not in $S(t)$. Then the interaction $\mathcal{H}_{-t}$ generated by integrating backwards cannot lie in S. If $\mathcal{H}_{-t}$ is not in S, the reason is likely to be that $\mathcal{H}_{-t}$ has long range interactions on a dimensionless scale. In fact, along a random trajectory integrated backwards one would reasonably expect all the interactions $\mathcal{H}_{-t}$ to have the same range in physical units. If this is the case the range of interaction must grow like e^t in dimensionless units for the interactions $\mathcal{H}_{-t}$. Hence, for large t, $\mathcal{H}_{-t}$ will generally lie outside S.

Trajectories and fixed points

The renormalization group equation (12.1) defines trajectories in the space S. On a trajectory, $\xi_t = e^{-t}\xi_0$. Thus, if the initial interaction $\mathcal{H}_0$ has a finite correlation length ξ_0, then the trajectory passes through all the surfaces of fixed ξ' with $\xi' < \xi_0$. This is illustrated by curve A in fig. 12.3a. If the initial interaction $\mathcal{H}_0$ has an infinite ξ_0, then the trajectory lies entirely on the critical surface S_c (curve B in fig. 12.3a).

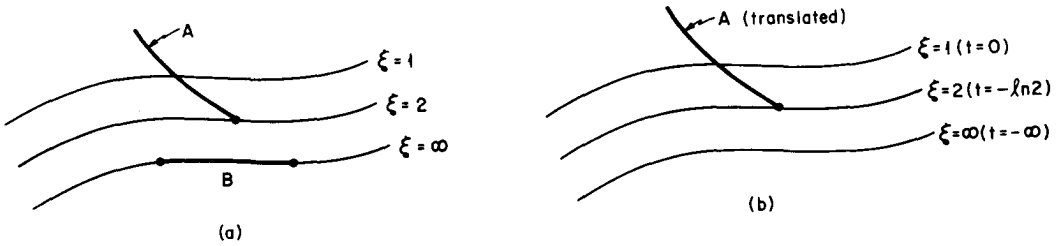

Fig. 12.3. (a) Two renormalization group trajectories A and B. Curve A begins at $\xi = 2$ and crosses lines of smaller ξ. Curve B begins on the critical surface ($\xi = \infty$) and stays there. (b) Trajectory A translated so that $\xi_0 = 1$.

Since the renormalization group transformation is independent of t, it is possible to translate trajectories in the variable t (as already seen in section 7). The only effect of a translation,

$$\mathcal{H}_t \to \mathcal{H}'_t = \mathcal{H}_{t-t_0} \tag{12.5}$$

is that the translated trajectory begins at t_0 instead of $t = 0$. One can use the freedom of translation in t to specify a normalization condition on ξ_0, say $\xi_0 = 1$. Then, $\xi_t = e^{-t}$. With this normalization the surfaces of equal ξ become surfaces of equal t for translated trajectories (fig. 12.3b).

A fixed point is an interaction $\mathcal{H}^*$ satisfying $U[\mathcal{H}^*] = 0$. A fixed point is a trivial trajectory: $\mathcal{H}_t = \mathcal{H}^*$ for all t solves (12.1). A fixed point must have a correlation length $\xi^* = 0$ or $\xi^* = \infty$ (since if $\mathcal{H}_t = \mathcal{H}^*$, then $\xi_t = \xi^*$, but, being a solution of eq. (12.1) means $\xi_t = e^{-t}\xi_0 = e^{-t}\xi^*$).

In order to illustrate the topological ideas, we will now develop the simplest form the topology can take. There are many ways in which the topology can be more complicated than described below, but complications can be understood once the simplest picture has been grasped.

Assume there are only two fixed points, one with $\xi^* = 0$, one with $\xi^* = \infty$. These fixed points are denoted P_0 and P_∞ in fig. 12.4. The fixed point P_∞ will be assumed to have the same properties as the fixed point (r^*, u^*) of section 4.

The simplest behavior one can obtain for a trajectory in the limit $t \to \infty$ is that it approaches a fixed point. This is not the only behavior possible in principle, but it is the only limiting behavior seen in previous sections and will be assumed here. Other possibilities are discussed in [150] and [154]. If all trajectories approach a fixed point, then the trajectories with finite ξ approach P_0 for $t \to \infty$ (e.g., curve A in fig. 12.4) and the trajectories on the critical surface approach P_∞ (curve B, fig. 12.4).

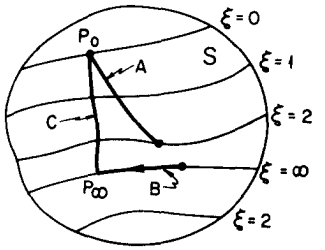

Fig. 12.4. Simple topology: the space S has two fixed points P_0 and P_∞. P_∞ lies on the critical surface and has $\xi = \infty$. P_0 has $\xi = 0$. Curve B runs into P_∞ and A runs into P_0.

There is a simple analogy in classical physics which provides an example of the topology being discussed [155]. Imagine a ball rolling on a hill according to the equation of motion

$$dx/dt = -\nabla V(x) \tag{12.6}$$

where $x = (x, y)$ are the horizontal coordinates of the ball and $V(x)$ is the elevation of the hill at the point x. (This equation is a simplification of Newton's law in that velocity (dx/dt) replaces mass times acceleration.) An example of a contour map of V is shown in fig. 12.5. It is assumed that $V(x)$ is analytic in x. The equation of motion is an example of (12.1); a fixed point is a stationary point of V. Two fixed points are shown in fig. 12.5; P_0 is an absolute minimum and P_∞ is a saddle point. If the ball is placed anywhere to the west of the ridge line R, the ball rolls to the bottom of the hill (P_0) and stops (the curve A in fig. 12.5 is an example of a trajectory of the ball). If the ball is placed exactly on the ridge line R, it rolls down to the saddle point P_∞ and stops (curve E, fig. 12.5). The ridge line R is analogous to the critical surface for the renormalization group equation (12.1).

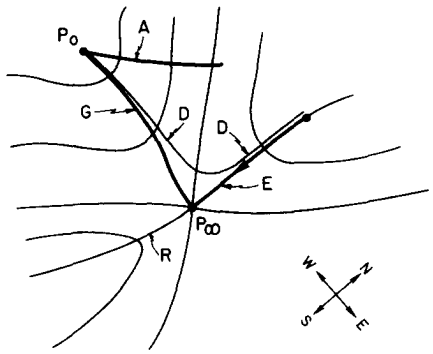

Fig. 12.5. Contour map realizing the topology of fig. 12.4. P_∞ is a saddle-point on the ridge R and P_0 is the bottom of the gully G. Path D shows the course of a ball initially near the ridge line, spending a great deal of time near the saddle point P_∞ and eventually rolling down the gully G and stopping at the bottom P_0.

It is interesting to study the trajectory of the ball when it is placed infinitesimally close to the ridge line R. Then it rolls down close to the ridge line almost to the saddle point. This is illustrated

by the curve D in fig. 12.5. Near the saddle point the ball veers away from the ridge line and rolls down the gully G. Near the saddle point the ball rolls very slowly because the terrain is almost flat.

Now let the initial location of the ball approach the ridge R. In this limit the curve D approaches a limiting curve E. The curve E has two parts; one part is a section of the ridge line R. The other part is the line G defining the bottom of the gully. The gully line G is also the trajectory of a ball which starts infinitesimally close to P_∞.

Consider again the simple example of the renormalization group with the two fixed points (fig. 12.4). Consider a trajectory which begins infinitesimally close to the critical surface (curve D in fig. 12.6). Assume the analogy to the classical equation of motion is valid. Then D has two parts; one part is essentially the trajectory E going to the fixed point P_∞. The second part is a trajectory G connecting the two fixed points. (The approximation of replacing D by G is equivalent to the approximation (4.41) of section 4. A change in t in the current lecture is equivalent to changing n in (4.41). The linear approximation of section 4 is appropriate only near P_∞ where the trajectory G is a straight line: changing either n or $T - T_c$ only changes one's location on G. The analogue in section 4 of D is obtained by keeping terms proportional to λ_2^n (λ_2 was defined in section 4) in eq. (4.41). The analogue in section 4 of E is then obtained by setting $T = T_c$ in eq. (4.41) which leaves only the λ_2^n terms.)

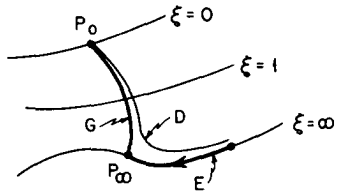

Fig. 12.6. Renormalization group trajectories near two fixed points. The curve D begins close to the critical surface. As the origin of D approaches the critical surface, the trajectory D approaches the trajectories E and G. For this example G is $S(\infty)$.

Now look at the space $S(\infty)$ defined earlier. With some assumptions, $S(\infty)$ is the curve G, including the two endpoints P_∞ and P_0. Starting at any interaction $\mathcal{H}$ on the curve G, one can extrapolate back along the trajectory an arbitrary amount of time without leaving the space S. The reason for this is that $\xi \to \infty$ on G as one approaches P_∞; to reach the point P_∞ with $\xi = \infty$ from any point with finite ξ requires an infinite amount of time (since $\xi_t = e^{-t}\xi_0$).

This means that any point on G is in $S(\infty)$. Starting from any other interaction in S not on G, and extrapolating back in t can take one outside the space S. For example, extrapolating the trajectory D of fig. 12.6 backwards would leave the figure. If so, $S(\infty)$ is only the curve G.

This completes the formulation of the simplest topology of the renormalization group. There are some general observations to be made. First of all, the primary role of the fixed points is to be "time sinks". Trajectories which go into a fixed point spend an infinite amount of time near the fixed point (just as the ball rolling down to the saddle point takes an infinite amount of time to reach the saddle point). This occurs because the velocity $d\mathcal{H}_t/dt$ goes to zero as $\mathcal{H}_t$ approaches the fixed point. Trajectories like D of fig. 12.6 which pass near a fixed point use up a lot of time near the fixed point for the same reason. The existence of these time sinks enables trajectories to spend an infinite amount of time in a compact region of S; otherwise there is danger of the trajectories leaving S, especially when one tries to move backwards in time along a trajectory.

The second observation is that in the example there is only one trajectory leaving the critical surface, namely, the trajectory G. The reason trajectories do not leave the critical surface anywhere other than P_∞ is that away from the fixed point the velocity $d\mathcal{H}_t/dt$ is non-zero; but where the velocity is non-zero it must point along the critical surface. As long as the velocity is continuous in S, the direction of the velocity has to be parallel to the critical surface for points infinitesimally close to the critical surface. Thus, trajectories like D stay close to the critical surface until they reach the vicinity of P_∞. At the point P_∞ there is no well-defined direction for the velocity; hence, the velocity can be perpendicular to the critical surface at points infinitesimally close to P_∞.

12.2. *Fixed points, subspaces, and renormalization*

Application to field theory

Now the problem of renormalization in quantum field theory will be discussed. The simple topology outlined above will be used to illustrate the general ideas. It was apparent from the original work of Gell-Mann and Low [149] that the problem of renormalization involves the renormalization group and its fixed points. The Gell-Mann–Low approach has been reviewed elsewhere [150]. What will be seen here is that a simple and general formulation of renormalized field theory can be given in terms of the space $S(\infty)$ defined earlier.

There are actually two parts to the renormalization problem. The first part is to obtain finite theories in the limit of an infinite cutoff. This part is very easy to discuss in terms of the space $S(\infty)$. The second part is to obtain theories with interaction. This will be discussed later in terms of Gaussian and non-Gaussian fixed points. It is the second part which is hard to achieve in four space-time dimensions, as will be seen in section 13. See however supplemental references.

The renormalization problem will be introduced using a conventional approach.

Suppose one is interested in solving a ϕ^4 interaction. To make the theory finite a cutoff Λ_0 is introduced (as in section 10). The problem is now to obtain a finite theory in the limit $\Lambda_0 \to \infty$.

An interaction with cutoff Λ_0 can be converted to an interaction with unit cutoff by a change of momentum scale, so that it becomes an interaction in S. There are two free parameters in the interaction (the bare mass and coupling constant) so the canonical surface is two-dimensional. The canonical surface is illustrated by the curve C in fig. 12.7.

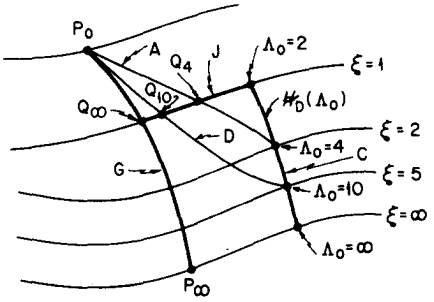

Fig. 12.7. Renormalization trajectories emanating from a canonical surface C of cutoff ϕ^4 theories. The canonical surface intersects the critical surface at $\Lambda_0 = \infty$. The trajectory G defines the renormalized theory. Trajectories such as G are referred to as "renormalized" trajectories.

Define a "canonical curve" $\mathcal{H}_D(\Lambda_0)$ to be a set of dimensionless interactions $\mathcal{H}_D$ in S, one for each value of Λ_0. If no renormalization is performed and the cutoff consists simply of a maximum momentum (rather than the lattice cutoff of section 10), then one finds the dimensionless parameters in $\mathcal{H}_D(\Lambda_0)$ to be $r_0(\Lambda_0) = \mu_0^2/\Lambda_0^2$ and $u_0(\Lambda_0) = \lambda_0 \Lambda_0^{d-4}$ where μ_0 and λ_0 are the bare mass and charge (r_0 and u_0 were defined in section 4). Define $\xi_D(\Lambda_0)$ to be the dimensionless correlation length for $\mathcal{H}_D(\Lambda_0)$. It was observed in section 10 that $\xi_D(\infty)$ must be ∞ in order that μ_R be finite and this is unlikely if no renormalization is performed.

Performing a renormalization conventionally means giving μ_0 and λ_0 a Λ_0 dependence such that the renormalized mass μ_R and renormalized coupling constant are cutoff independent. In the simple topology considered here it will turn out that it is sufficient to keep μ_R cutoff-independent (see below). A procedure for accomplishing this has already been suggested in section 10. If one is given the dimensionless parameters r_0 and u_0, the dimensionless correlation length ξ is determined independently of Λ_0: $\xi = \xi(r_0, u_0)$. Next specify the desired value of μ_R. Now *choose* Λ_0 to be $\Lambda_0 = \mu_R\, \xi(r_0, u_0)$. (This is analogous to choosing the lattice spacing a to be $1/\mu_R\, \xi(b, u_0, K)$ in section 10.) To obtain a curve $\mathcal{H}_D(\Lambda_0)$ suppose one chooses a fixed value for u_0. Then one can solve the equation $\Lambda_0 = \mu_R\, \xi(r_0, u_0)$ to give r_0 as a function of Λ_0. Knowing $r_0(\Lambda_0)$ one determines $\mu_0^2 = r_0(\Lambda_0)\Lambda_0^2$ and $\lambda_0 = u_0 \Lambda_0^{4-d}$. Now $\Lambda_0 \to \infty$ corresponds to $r_0 \to r_{0c}(u_0)$ where $r_{0c}(u_0)$ is the value of r_0 on the critical surface. The renormalized mass (in physical units) is $\Lambda_0/\xi(r_0, u_0)$ (see section 10) and therefore μ_R is independently of r_0 by construction.

The curve $\mathcal{H}_D(\Lambda_0)$ is illustrated in fig. 12.7. In fig. 12.7 the canonical surface is replaced by a curve (otherwise one would have more than two dimensions in the illustration which is impractical). The curve $\mathcal{H}_D(\Lambda_0)$ is a piece of the curve C in fig. 12.7. It will now be shown that this canonical curve defines a renormalized theory which is independent of u_0.

Suppose one solves the renormalization group equation (12.1) with $\mathcal{H}_D(\Lambda_0)$ as the initial interaction. This gives a set of curves $\mathcal{H}_t(\Lambda_0)$ in the space S. For illustration the curve A in fig. 12.7 is $\mathcal{H}_t(4)$, the curve D is $\mathcal{H}_t(10)$. Since the scale factor for rescaling $\mathcal{H}_D(\Lambda_0)$ has already been specified, namely Λ_0, we know the scale factor for rescaling $\mathcal{H}_t(\Lambda_0)$, namely $e^{-t}\Lambda_0$.

Now observe the following. For each Λ_0 there is a value of t for which $\mathcal{H}_t(\Lambda_0)$ intersects the surface of interactions with $\xi = 1$. The set of such intersections defines the curve J shown in fig. 12.7. For example, the trajectory $\mathcal{H}_t(4)$ (A in fig. 12.7) intersects the curve J at the point Q_4 in fig. 12.7. The interaction Q_4 defines the same physics as the interaction $\mathcal{H}_D(4)$. Likewise the interaction Q_{10} defines the same physics as the interaction $\mathcal{H}_D(10)$. The only change is a scale transformation. It is easily seen that the scale transformation is the same for both Q_4 and Q_{10}. The reason is that $\mathcal{H}_D(\Lambda_0)$ was defined to have a constant correlation length in physical units (namely μ_R^{-1}) independent of Λ_0; since Q_4 and Q_{10} define the same physics as $\mathcal{H}_D(4)$ and $\mathcal{H}_D(10)$ respectively, they also generate the same physical correlation length. Since Q_4 and Q_{10} have the same dimensionless correlation length 1, the rescaling factor is the same for both (namely, momenta are multiplied by $\mu_R/\xi = \mu_R$).

Now let $\Lambda_0 \to \infty$. In this limit Q_4 and Q_{10} are replaced by Q_∞, the intersection of J with the curve G. Q_∞ is a well-defined, cutoff interaction so it describes a well-defined physics including a complete set of vacuum expectation values. Hence, the limit $\Lambda_0 \to \infty$ exists: there is a finite (renormalized) theory in the limit $\Lambda_0 \to \infty$.

In this formulation of renormalization one has mass renormalization (because μ_0 depends on Λ_0). There is only a trivial coupling constant renormalization (u_0 is held fixed instead of λ_0) and the renormalized theory does not depend on u_0. There is no free coupling constant in the renormalized theory; the only free parameter in the renormalized theory is μ_R. The reason for this is the simple topology assumed here; in a later section there will be examples of renormalized theories with more free parameters. The renormalized theory associated specifically with the

nontrivial fixed point r^*, u^* of section 4 is an example where the renormalized coupling constant is fixed [156]. This is evident from sections 7 and 8; the expressions there for long range correlation functions depend only on ξ or r respectively. It was already shown in [150] that there could be fixed points where the renormalized coupling constant is fixed and the unrenormalized coupling constant is arbitrary (these were called "infra-red stable" in [150]).

There is also wavefunction renormalization in the theory. This is because the renormalized field $\phi_R(x)$ can differ at most by a finite scale factor from the Fourier transform of the spin variable σ'_q in Q_∞. This spin variable will differ by an infinite scale factor from the original spin variable σ_q in $\mathcal{H}_D(\infty)$ because the trajectory from $\mathcal{H}_D(\infty)$ to Q_∞ spends an infinite amount of time at the fixed point P_∞. (The original spin variable σ_q is likely to differ by an infinite scale factor from the unrenormalized field: see, e.g., eq. (10.46) with $a = 0$ to correspond to $\Lambda_0 = \infty$; it is unlikely the two infinite scale factors combine to make a finite factor.)

The fact that the renormalized theory does not depend on u_0 is a field theoretic example of universality. One has even greater freedom: one can choose the canonical curve $\mathcal{H}_D(\Lambda_0)$ to be the trajectory G itself and still obtain the same renormalized theory. (This assumes one is willing to consider curves $\mathcal{H}_D(\Lambda_0)$ not on the original canonical surface.)

The generalization of this result is the following. Any trajectory of the renormalization group equation (12.1) which (like G) can be extrapolated back to the critical surface ($t = -\infty$) without leaving the space S defines a renormalized theory. The reason is simple: one can choose the bare interaction $\mathcal{H}_D(\Lambda_0)$ to be the point on the trajectory with $\xi = \Lambda_0/\mu_R$. As long as the trajectory extrapolates back to the critical surface, it has points with arbitrarily large ξ, so one can let $\Lambda_0 \to \infty$ without difficulty. Any such trajectory will be called a "renormalized trajectory". An example of a renormalized trajectory analogous to G was discussed in section 7, and in part in section 4 (as explained above). Conversely, any trajectory which, when extrapolated backwards, leaves the space S before reaching the critical surface, defines a nonlocal theory of no interest. There is no guarantee that a given renormalized trajectory can be reached starting from a given canonical surface. In the example the trajectories leaving the canonical surface (A and D in fig. 12.7) approach G because that is the only renormalized trajectory in the example.

The subspace $S(\infty)$ is the same as the set of all renormalized trajectories; thus, the set $S(\infty)$ is simply the set of all possible renormalized interactions, in the sense that one can think of Q_∞ as a renormalized interaction.

12.3. *Multiple fixed points, domains, and universality*

Several fixed points; domains

In practical situations there are usually more than two fixed points of the renormalization group. The next topic is to discuss some of the complications that arise when there are several fixed points. To illustrate this problem imagine that there are three fixed points $P_{A\infty}$, $P_{B\infty}$, and $P_{C\infty}$ on the critical surface (fig. 12.8) in addition to P_0. It will still be assumed that all trajectories go to a fixed point as $t \to \infty$.

Fixed points can be classified according to their stability or instability. A fixed point is stable if all trajectories in the neighbourhood approach the fixed point. P_0 is the only stable fixed point in the example. Other fixed points are unstable. There are, however, degrees of instability. In the example it is assumed that trajectories near $P_{A\infty}$ on the critical surface (for example B in fig. 12.8) approach $P_{A\infty}$ while trajectories off the critical surface go away from $P_{A\infty}$. The point $P_{A\infty}$ is called once unstable. Likewise, the fixed point $P_{C\infty}$ is once unstable, according to fig. 12.8. The fixed point $P_{B\infty}$ is twice unstable: trajectories on or off the critical surface move away from $P_{B\infty}$. This is because (assuming there are no other fixed points) the velocity $d\mathcal{H}_t/dt$ must always point towards

$P_{A\infty}$ on the critical surface between $P_{A\infty}$ and $P_{B\infty}$; the velocity cannot change direction without going through zero.

The above discussion takes literally fig. 12.8, in which the space S is two-dimensional and the critical surface is one-dimensional. The general definition of n-unstable fixed points for arbitrary n in an infinite-dimensional space S will be given later.

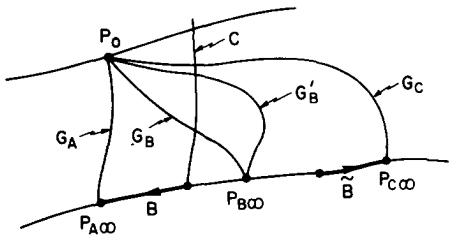

Fig. 12.8. Topology of the renormalization group in a region having three fixed points $P_{A\infty}$, $P_{B\infty}$ and $P_{C\infty}$ on the critical surface. $P_{A\infty}$ and $P_{C\infty}$ are once-unstable. $P_{B\infty}$ is twice-unstable.

There are unique renormalized trajectories G_A and G_C leaving the fixed points $P_{A\infty}$ and $P_{C\infty}$, analogous to the unique trajectory G in fig. 12.6 or the gully G in the classical analogue. In addition there is an infinite set of renormalized trajectories leaving the point $P_{B\infty}$. The point $P_{B\infty}$ is analogous to the top of a hill in the classical analogue while $P_{A\infty}$ and $P_{C\infty}$ are analogous to saddle points. The appropriate contour map is given in fig. 12.9. The curves G_B and G'_B in both figs. 12.8 and 12.9 are examples of trajectories leaving $P_{B\infty}$.

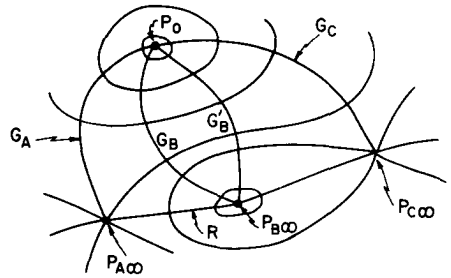

Fig. 12.9. Contour map of the topology of fig. 12.8.

When there are several fixed points an additional problem arises. Suppose one has a canonical surface (C in fig. 12.8). The question is which fixed point is connected via the renormalization group to the canonical surface. In fig. 12.8 the renormalization group trajectories leaving the canonical surface go to either P_0 or $P_{A\infty}$. In the case of an interaction on the critical surface the trajectory goes to $P_{A\infty}$. So, the critical behavior is governed by $P_{A\infty}$, not $P_{B\infty}$ or $P_{C\infty}$, and renormalization starting from the canonical surface is governed by the single renormalized trajectory G_A.

Associated with each fixed point $\mathcal{H}^*$ in S there is a domain $D(\mathcal{H}^*)$ of interactions in S. The domain $D(\mathcal{H}^*)$ is defined as the set of all initial interactions $\mathcal{H}_0 \in S$ such that the trajectory $\mathcal{H}_t$

starting from $\mathcal{H}_0$ approaches $\mathcal{H}^*$ as $t \to \infty$. In fig. 12.8 the domain of P_0 consists of all points above the critical surface; the domain of $P_{A\infty}$ consists of all points on the critical surface to the left of $P_{B\infty}$; the domain of $P_{B\infty}$ consists solely of the point P_B, etc. A fixed point $\mathcal{H}^*$ is connected to a given canonical surface C if the surface C intersects the domain $D(\mathcal{H}^*)$.

The dimensionality of the domain $D(\mathcal{H}^*)$ is connected to the degree of instability of $\mathcal{H}^*$: if $\mathcal{H}^*$ is an n-unstable fixed point then the domain $D(\mathcal{H}^*)$ has *n fewer* dimensions than S. This is easily seen in the example of fig. 12.8. It will be equally trivial to see in general once the definition of n-unstable has been given.

The dimensional rule for domains makes the highly unstable fixed points very elusive. Suppose the canonical surface is a curve (the only free parameter being a temperature or bare mass). Then the canonical surface is unlikely to intersect any domain for a twice-unstable fixed point like $P_{B\infty}$. This is evident in fig. 12.8. This rule holds provided that topological structures like fixed points, the critical surface, etc. are randomly located in the space S, relative to any *a priori* chosen canonical surface. An exception to this rule is in the case of symmetry; see later discussion.

In critical phenomena, critical points are classified by the number of thermodynamic parameters that are fixed at the critical point. An ordinary ferromagnetic critical point requires the temperature T and the magnetic field to be fixed, so the corresponding fixed point is twice unstable. If, as in these sections one defines the space S itself to exclude magnetic fields, then the only fixed parameter is T and the fixed point is once-unstable. Thus, in the example of fig. 12.8, the points $P_{A\infty}$ and $P_{C\infty}$ could each describe the critical behavior of an ordinary critical point. The existence of the two fixed points would mean the existence of two distinct sets of critical exponents; different substances (corresponding to different canonical surfaces) would separate into two classes, each class separately showing a common critical behavior. This is still the universality hypothesis but in a more restricted form than earlier formulations; see, e.g., [157].

There are critical points where three thermodynamic parameters are fixed instead of two. These are called tri-critical points; see, e.g., [158]. Excluding the magnetic field again, there are two fixed parameters. The extra thermodynamic parameter is, e.g., a concentration. The most popular tri-critical point is in ^{3}He–^{4}He mixtures: the two thermodynamic parameters are the ^{3}He concentration and the temperature. A tri-critical point corresponds to a twice-unstable fixed point like $P_{B\infty}$ (excluding external fields: in the ^{3}He–^{4}He system there is no way in practice to imitate the external field of magnetic systems). For dimensions near 4 the Gaussian fixed point of section 3 is twice-unstable; Riedel and Wegner [159] have proposed that this fixed point describes the ^{3}He–^{4}He tri-critical point.

In field theory, the degree of instability of a fixed point determines the number of free parameters in the renormalized theory associated with the fixed point. In the example discussed earlier (fig. 12.7) there is a once-unstable fixed point and only one free parameter, namely the renormalized mass μ_R. (The effective interaction for cutoff $\mu_R e^t$ is uniquely determined as the point on the trajectory G in fig. 12.7 with $\xi = e^{-t}$. Hence there are no other free parameters.) For a twice unstable fixed point there is a one parameter family of renormalized trajectories; in this case the parameter labelling these trajectories is a second parameter in the renormalized theory, in addition to μ_R.

Thus the renormalized field theory associated with the nontrivial fixed point of section 4, for $d < 4$, has no free coupling constants. In contrast, the fixed point discussed by Gell-Mann and Low [149, 150] has to correspond to a twice unstable fixed point in order that there be two free parameters (mass and charge) in the renormalized theory. (The once unstable fixed points discussed here correspond to the "infra-red stable" fixed points of [150] while "twice unstable" here corresponds to "ultraviolet stable" in [150].) When one produces a renormalized theory with two free parameters such as the conventionally renormalized ϕ^4 theory, the canonical

surface also has to be at least two-dimensional and the canonical curve $\mathcal{H}_D(\Lambda_0)$ involves two parameters depending on Λ_0 (e.g. $\mu_0(\Lambda_0)$, $\lambda_0(\Lambda_0)$). This will not be discussed in detail; however, see section 13. In conclusion, one needs both coupling constant renormalization (λ_0 depending on Λ_0) and a twice-unstable fixed point in order to have an arbitrary renormalized coupling constant in the renormalized theory.

There is an argument (see [150]; it is omitted here for lack of space) that only fixed points of limited instability will be relevant to elementary particles. The hypothesis, stated more precisely, is that only *stable* fixed points are relevant provided the space S is restricted to interactions conserving a suitably chosen internal symmetry. There can be instabilities which break the symmetry (instabilities here correspond to "generalized mass terms" in [150]) but hopefully not too many. This hypothesis is the basis for obtaining interactions with only a few free parameters; it substitutes for the canonical approach of considering only simple polynomial interactions.

Symmetries

The next topic is the effect of space-time or internal symmetries on the topological analysis. This is a subject that can become arbitrarily complicated; only a few comments will be made here; see also [160–162].

Essentially any symmetry that can be incorporated into an interaction $\mathcal{H}$ will be preserved by the renormalization group transformation U (as defined in section 11). U preserves rotational symmetry (Euclidean symmetry in the case of field theory). It preserves the symmetry for $\sigma_q \to -\sigma_q$ in the absence of a magnetic field. If σ_q has internal components (as in a Heisenberg ferromagnet or a field theory with isospin (see section 9)), U preserves internal rotational symmetry. When a symmetry is possible there will be a subspace of S of interactions which incorporate the symmetry. There are also interactions which break the symmetry: lattice models have cubic symmetry but not full rotational symmetry, interactions with external magnetic fields break the symmetry $\sigma_q \to -\sigma_q$.

The subspaces of S of interactions preserving a symmetry are invariant to U. If one is uninterested in interactions which break a symmetry, one can redefine S to be the subspace of interactions which preserve the symmetry and carry out the topological analysis of this lecture in the subspace. In fact, the space S of this section has been understood to omit interactions with an external field.

A consequence of this must be noted. Suppose one is interested in the critical point of a ferromagnet without an external field. As observed above, this critical point corresponds to a once unstable fixed point in a space S which preserves the symmetry $\sigma_q \to -\sigma_q$. In the larger space S' of all interactions the same fixed point could be highly unstable; there is no guarantee that it is just twice unstable, as claimed earlier. Further instabilities ought to imply the existence of further thermodynamic parameters besides the external field which are fixed at the critical point. However, setting the external field to zero means no interactions are present which break the spin reflection symmetry: in effect, setting the external field to zero also sets the extra thermodynamic parameters equal to zero too.

It may happen that all the fixed points in S preserve a symmetry, say rotational symmetry. Then, even if the initial interaction violates the symmetry, the renormalized theory associated with the fixed points preserve the symmetry. For example, in fig. 12.7 the curve C might represent a set of nearest neighbor interactions on a lattice with only cubic or hypercubic symmetry. The lattice is used to ensure a positive metric; see section 10. If the curve G lies in the subspace of complete Euclidean symmetry, then the renormalized theory will be Euclidean invariant. In order that G lie in the Euclidean-invariant subspace S_R, it is necessary that the fixed points P_∞ and P_0 and the trajectory G still be present when S is replaced by S_R. It is of course crucial that U itself

preserves Euclidean symmetry: otherwise a trajectory like G is unlikely to stay inside S_R.

There is no *a priori* guarantee that all fixed points are in S_R or that the trajectories like G leaving a fixed point all remain inside S_R. More generally, to achieve Euclidean invariance and positivity simultaneously one needs to have trajectories originating with a nearest neighbor interaction and ending in S_R. Whether this is the case has to be checked for each situation of interest.

It seems likely that this condition can be satisfied within perturbation theory (small u) but this has not been carefully checked.

There can be fixed points which violate possible symmetries. A specific example with some fixed points violating an internal rotational symmetry and others preserving it has been discussed elsewhere [160]. The least unstable fixed points may preserve or may violate the symmetry; only a calculation can determine which. If a fixed point violates a continuous symmetry, one can use the symmetry transformations to construct a fixed surface from the fixed point.

Consider again the problem of renormalization. As found in section 3, there is a Gaussian fixed point; there is also a curve like G in fig. 12.6 which contains only Gaussian interactions. In less than four (space-time) dimensions there is also a non-Gaussian fixed point, but in four dimensions the Gaussian fixed point is the only one known so far (see section 13). This suggests that even if the canonical surface contains non-Gaussian interactions, the renormalized theory will be the purely Gaussian theory defined by G, i.e., a free field theory. The ability of a canonical surface of interacting theories to produce free field theories after renormalization is closely related to coupling constant renormalization in perturbation theory; see, e.g., section 3 G of [150].

Thus, the serious problem of renormalization theory in $d = 4$ is to find renormalized theories with interaction. The simplest way to ensure the existence of such theories would be to find a non-Gaussian fixed point, or, at least a finite boundary for the domain of the Gaussian fixed point; see section 13.

12.4. *Fixed points and anomalous dimensions*

Linearized renormalization group equations

The next topic is to discuss the renormalization group transformation near a fixed point. Linearized equations near the fixed point will be defined. The eigenfunctions of the linearized equations define a complete basis set of local operators (in statistical mechanics these are Kadanoff's local operators [157, 163]). The eigenvalues are the anomalous dimensions (Kadanoff's scaling indices) of the local operators. Operators with anomalous dimension less than d are called "relevant". The number of such operators is the degree of instability of the fixed point. Operators with dimension $>d$ are irrelevant. Operators with dimension $=d$ are a nuisance (they may or may not contribute to the degree of instability; they give rise to long-lived transients (see section 7) and other complications [152]).

Wegner [152] gives a discussion of behavior near a fixed point going beyond the linear approximation.

The analysis will continue to be abstract with general assumptions that have been confirmed only by examples.

Consider a trajectory $\mathcal{H}_t$ satisfying (12.1) and suppose that $\mathcal{H}_t \approx \mathcal{H}^*$ where $\mathcal{H}^*$ is a fixed point. Then one can write

$$\mathcal{H}_t = \mathcal{H}^* + \delta\mathcal{H}_t \tag{12.7}$$

where $\delta\mathcal{H}_t$ is small. The perturbation $\delta\mathcal{H}_t$ satisfies an equation

$$\partial\delta\mathcal{H}_t/\partial t = L \cdot \delta\mathcal{H}_t + O[\delta\mathcal{H}_t^2] \tag{12.8}$$

where $O[\delta\mathcal{H}_t^2]$ refers to terms of second order in the coupling constants of $\delta\mathcal{H}_t$. The operator L is a linear operator; the operator L associated with the transformation of section 11 is given in the Appendix. L depends on the fixed point $\mathcal{H}^*$. The 2 x 2 matrix M of eq. (4.32) is also an example of an operator L.

The initial interaction $\delta\mathcal{H}_0$ is normally the integral over all space of an energy density. It will however be convenient to consider localized perturbations $\delta\mathcal{H}_0$ also. To be general write,

$$\delta\mathcal{H}_0 = \int_x g(x) O[x; \sigma] \tag{12.9}$$

where $g(x)$ is an arbitrary small function and $O[x; \sigma]$ is an arbitrary interaction density.

It is convenient to consider translation-invariant densities $O[x; \sigma]$. Examples of translationally invariant densities $O[x; \sigma]$ are $s(x)$; $\int_y \exp\{-(x-y)^2\} s(y)$; $\int_y \int_z \{\exp -(x-y)^2\} \exp\{-(x-z)^2\} \times s(y)s(z)$, etc., where $s(x)$ is the Fourier transform of σ_q. (For a formal definition, see the Appendix.) In the linear approximation $\delta\mathcal{H}_t$ is linear in $g(x)$. One can write,

$$\delta\mathcal{H}_t = \int_x g(x)\, O[x e^{-t}; \sigma; t] \tag{12.10}$$

where the argument of O is written $x e^{-t}$ instead of x for reasons to be explained shortly.

It is natural to look for exponential solutions of (12.8); namely, solutions

$$\delta\mathcal{H}_t = \exp(-d_m t)\, O_m[\sigma]. \tag{12.11}$$

The functionals O_m must satisfy an eigenvalue equation:

$$-d_m O_m[\sigma] = L \cdot O_m[\sigma]. \tag{12.12}$$

Explicit examples of solutions in terms of localized functionals $O_m[\sigma]$ are given in the Appendix. The operator L is not Hermitian, which means there might also be solutions behaving like $t^k \exp(-d_m t)$ but these have not been encountered in practice and will be ignored. We assume that the set of solutions $O_m[\sigma]$ are complete; i.e. any localized functional can be expressed as a linear combination of the $O_m[\sigma]$ (see later for a more accurate expansion theorem).

The functionals $O_m[\sigma]$ are localized about the origin. One can define functionals $O_m[x; \sigma]$ centered about any point x by a translation. The rule (see the Appendix) is

$$O_m[x; \sigma] = O_m[\sigma'] \tag{12.13}$$

where $\sigma'_q = \exp(iq \cdot x)\, \sigma_q$. These translated densities also define solutions of the linearized equation, in the form

$$\delta\mathcal{H}_t = \exp(-d_m t)\, O_m[x e^{-t}; \sigma]. \tag{12.14}$$

This is proven in the Appendix. The reason for the argument $x e^{-t}$ is that the physical location of an interaction must be independent of t. The factor e^{-t} compensates for the changing dimensionless length scale.

Now suppose one has an arbitrary initial interaction of the form (12.9) where $O[x; \sigma]$ is an arbitrary translated interaction. The operator $O[0; \sigma]$ can be expressed as a linear combination of the $O_m[\sigma]$ (by assumption):

$$O[0; \sigma] = \sum_m c_m O_m[\sigma]. \tag{12.15}$$

This immediately implies that

$$O[x;\sigma] = \sum_m c_m O_m[x;\sigma]. \tag{12.16}$$

Therefore the solution $\delta\mathcal{H}_t$ corresponding to the initial condition is,

$$\delta\mathcal{H}_t = \int_x g(x) \sum_m c_m \exp(-d_m t)\, O_m[x e^{-t};\sigma]. \tag{12.17}$$

Warning: in practice the expansion (12.17) is not always convergent. What does seem to be true is that asymptotically for large t the expansion (12.17) is valid. An abstract example that illustrates this is the following. Let $f(q)$ be a function like $(1+q^2)^{-1}$ whose Taylor series in q^2 converges only for $q^2 < 1$. Consider the function $f(qe^{-t})$ for given q with $q^2 > 1$. For $t = 0$ the Taylor series expansion in q^2 diverges. For sufficiently large t the series is valid.

In statistical mechanics, localized perturbations arise due to local fluctuations in temperature, chemical potential, etc.; each such fluctuation corresponds to a particular functional $O[x;\sigma]$. In field theory these functionals correspond to local composite operators, after renormalization. (At first sight the typical localized operators like $\int_x \int_y \exp(-x^2 - y^2) s(x) s(y)$ do not look like the strictly local products like $\phi^2(x)$ of a field theory. However, a dimensionless separation $|x - v| \sim 1$ in $\delta\mathcal{H}_0$ becomes a separation Λ_0^{-1} in physical units, and when $\Lambda_0 \to \infty$ the physical separation goes to 0.) The specific functionals $O_m[x;\sigma]$ correspond to Kadanoff's local operators [157, 163] with scaling indices d_m. In field theory they define a basis of local scale invariant operators $O_m(x)$ [164]; d_m is the anomalous dimension of O_m. (See below for further discussion.)

So far the analysis has assumed that $\mathcal{H}_0 \approx \mathcal{H}^*$ so that $\delta\mathcal{H}_0$ is small. For a more general analysis suppose $\mathcal{H}_0^c$ is the critical interaction of some system, on the critical surface in the domain of $\mathcal{H}^*$ but not near $\mathcal{H}^*$. Then there is a critical trajectory $\mathcal{H}_t^c$ which approaches $\mathcal{H}^*$ for $t \to \infty$. Consider small departures from the critical trajectory

$$\mathcal{H}_t = \mathcal{H}_t^c + \delta\mathcal{H}_t. \tag{12.18}$$

In this section $\delta\mathcal{H}_t$ is assumed to be small which means (in the language of section 7) that t is either in the initial transient region where $\mathcal{H}_t^c$ is appreciably different from $\mathcal{H}^*$ or in an intermediate range of t with $\mathcal{H}_t^c \approx \mathcal{H}^*$ and $\delta\mathcal{H}_t$ still small. The "correlation length region" for very large t where $\delta\mathcal{H}_t$ is large will be ignored here. In the transient region $\delta\mathcal{H}_t$ satisfies a linearized equation but with a t-dependent operator $L[\mathcal{H}_t^c]$. For larger t where $\mathcal{H}_t^c \approx \mathcal{H}^*$, $L[\mathcal{H}_t^c]$ reduces to the operator L defined previously. There is an expansion theorem for large t: $\delta\mathcal{H}_t$ satisfies (12.8) for large t and therefore has an expansion of the form (12.17) with coefficients c_m independent of t.

To show that d_m is an anomalous dimension it is useful to consider correlation functions involving a localized perturbation $O[x;\sigma]$. One can calculate correlation functions such as $Z^{-1}\langle s(x_1) \ldots s(x_m)\, O[x;\sigma] \exp(\mathcal{H}_0^c)\rangle$ from the generating functional

$$Z(j;g) = Z^{-1} \int_\sigma \exp\left\{ \int_x j(x) s(x) + \int_x g(x)\, O[x;\sigma] + \mathcal{H}_0^c[\sigma] \right\} \tag{12.19}$$

with the g term treated to first order. Now define the initial interaction $\mathcal{H}_0$ to be $\mathcal{H}_0^c + \int_x g(x) \times O[x;\sigma]$. Since g can be treated to first order $\mathcal{H}_t$ has the form, implied by (12.7):

$$\mathcal{H}_t = \mathcal{H}^* + \int_x g(x) \sum_m c_m \exp(-d_m t)\, O_m[x e^{-t};\sigma]. \tag{12.20}$$

(If one is using the renormalization group equation of section 11, the generating functional (12.19) is obtained by substituting $\mathcal{H}_t$ into eq. (11.32) of section 11.)

One is now led to consider correlation functions involving the local operators $O_m[x;\sigma]$ and calculated using $\mathcal{H}^*$, for example

$$f_m^*(x; x_1, \ldots, x_n) = (Z^*)^{-1} \langle s(x_1) \ldots s(x_n)\, O_m[x;\sigma]\, \exp\{\mathcal{H}^*[\sigma]\}\rangle \tag{12.21}$$

where Z^* is the partition function for $\mathcal{H}^*$. This correlation function can be shown to satisfy a scaling law analogous to eq. (7.45) of section 7:

$$f_m^*(x; x_1, \ldots, x_n) = \exp(-nd_s t)\exp(-d_m t)\, f_m^*(x\mathrm{e}^{-t}; x_1\mathrm{e}^{-t}, \ldots, x_n\mathrm{e}^{-t}) \tag{12.22}$$

which means d_m is the anomalous dimension of O_m just as d_s is the anomalous dimension of the spin s. The proof of (12.22) will only be sketched. One can compute $f_m^*(x; x_1, \ldots, x_n)$ from the generating functional

$$(Z^*)^{-1}\langle s(x_1) \ldots s(x_n) \exp\left\{\mathcal{H}^*[\sigma] + \int_x g(x) O_m[x;\sigma]\right\}\rangle \tag{12.23}$$

with the g term treated to first order. The renormalization group equations can be integrated with $\mathcal{H}^* + \int_x g(x) O_m[x;\sigma] = \mathcal{H}_0$. The result is that the above generating functional can be written

$$(Z^*)^{-1} \exp(-nd_s t)\langle s(x_1\mathrm{e}^{-t}) \ldots s(x_n\mathrm{e}^{-t}) \exp\left\{\mathcal{H}^*[\sigma] + \int_x g(x)\exp(-d_m t) O_m[x\mathrm{e}^{-t};\sigma]\right\}\rangle \tag{12.24}$$

from which the result follows. [If one is working with the explicit form of the renormalization group equations of section 11, the expression,

$$\langle s(x_1) \ldots s(x_m) \exp(\mathcal{H}_t)\rangle \tag{12.25}$$

must be understood as a shorthand for

$$\frac{\delta}{\delta j(x_1, t)} \cdots \frac{\delta}{\delta j(x_m, t)}\, Z[j, \mathcal{H}_t] \tag{12.26}$$

where $Z[j, \mathcal{H}_t]$ is given by (11.32) and

$$j(x, t) = \int_q \exp(iq \cdot x) j(q, t). \tag{12.27}$$

Furthermore, one should integrate from t' to $t' + t$ say with t' large instead of 0 to t in order to be able to use the prescription (12.26) throughout the range of t. Also, the function $\rho(t)$ is assumed to be bt where b is a constant (see the Appendix); d_s is given by $\frac{1}{2}d - b$. The function $\rho(t)$ is constructed not to change due to the perturbation $\int g(x) O[x, \sigma]$. This is possible because the function $\rho(t)$ was needed only to ensure the existence of the fixed point $\mathcal{H}^*$; there is no need to modify $\rho(t)$ when considering perturbations about $\mathcal{H}^*$ (except in the case that one is looking for nearby fixed points).]

A consequence of the scaling law for correlation functions involving $O_m[x;\sigma]$ is that the operators with the smallest values of d_m have the largest long range correlation functions. Long range means $|x_i - x_j| \gg 1$ in the dimensionless units of the original interaction, i.e. long range compared to the lattice spacing in statistical mechanics, long compared to the cutoff distance Λ_0^{-1} in field theory. Thus, when one can write the expansion (12.20) the operator $O[x;\sigma]$ can be replaced by the leading operator $O_m[x;\sigma]$ in its expansion (i.e. the operator with lowest d_m for

which c_m is nonzero) in any long range correlation function. This means in the simple model of section 4 that almost all operators $O[x;\sigma]$ odd in σ behave like $s(x)$ itself and almost all operators even in σ behave like the energy density (generated by fluctuations in temperature T). This is part of Kadanoff's operator theory [157, 163].

Kadanoff's operator reduction formula [163] (the operator product expansion in field theory [164, 165]) can be derived within the framework described here. No details will be reported here. The idea is that if one is interested in the product $O_{m_1}(x_1;\sigma)\,O_{m_2}(x_2;\sigma)$ with $|x_1 - x_2|$ large, inside a correlation function with spins $s(x_3)$, $s(x_4)$, etc. very far away ($|x_3 - x_1| \gg |x_2 - x_1|$, etc.) then one can put the product into $\mathcal{H}_0$ by defining

$$\mathcal{H}_0 = \mathcal{H}^* + \int_{x_1}\int_{x_2} g(x_1, x_2)\, O_{m_1}(x_1;\sigma)\, O_{m_2}(x_2;\sigma). \tag{12.28}$$

This interaction is now treated to first order in g; for large t one will have

$$\mathcal{H}_t = \mathcal{H}^* + \int_{x_1}\int_{x_2} g(x_1, x_2)\, c_m(x_1 - x_2)\, \exp(-d_m t)\, O_m(x_1 e^{-t};\sigma) \tag{12.29}$$

which replaces the product of two operators by a sum over the basis.

The next topic is translationally invariant perturbations [152]. This means $\delta\mathcal{H}_0$ has the form (12.9) with $g(x) = g$ independent of x. In this case one makes a change of variables $x \to xe^t$ in (12.17) giving

$$\delta\mathcal{H}_t = g \sum_m c_m \exp\{(d - d_m)t\} \int_x O_m[x;\sigma]. \tag{12.30}$$

In this expression the terms with $d_m < d$ increase with t while those with $d_m > d$ decrease with t. Hence the number of relevant operators O_m ($d_m < d$) determines the number of instabilities of the fixed point $\mathcal{H}^*$. Thus one defines an n-unstable fixed point $\mathcal{H}^*$ as a fixed point with n relevant operators. (If there are any nuisance operators ($d_m = d$) a second order calculation in g (see [152] and section 13) is required to determine whether they count as instabilities or not.)

The part of the domain of $\mathcal{H}^*$ near $\mathcal{H}^*$ itself is easily determined; it consists of all perturbations not containing any relevant operators. This means the domain has n less dimensions than S.

13. Futile (so far) search for a non-trivial fixed point – ϕ^4 field theory in 4 dimensions

It was observed in section 12 that for any fixed point $\mathcal{H}^*$, there is a domain D of interactions whose critical behavior is described by $\mathcal{H}^*$. It is of crucial importance both to statistical mechanics and field theory to be able to locate boundaries of such domains.

An approximate technique for locating a domain boundary will be described here using high temperature expansions. The technique will be developed for a particular example, namely the problem of locating a boundary to the domain of the Gaussian fixed point in four dimensions. (This domain will be called the "free field domain".) The method, with further refinement, should be applicable to other problems involving domain boundaries.

The nature of the free field domain in four dimensions is of special interest to quantum field theorists. Consider the problem of renormalizing the ϕ^4 field theory. As long as the ϕ^4 interaction, suitably cutoff, lies in the free field domain, the renormalized field theory defined as in section 12 will have no interaction. The free field domain certainly includes the case of small bare coupling

constants (see later); the question is whether for *large* coupling constants the interaction lies outside the free field domain and therefore perhaps in the domain of a nontrivial fixed point.

It is perfectly possible that the entire critical surface (see section 12) is within the free field domain. Then there would be no way to obtain a renormalized ϕ^4 theory with interaction. This does not conflict with the existence of a renormalized perturbation theory for the ϕ^4 interaction. One already knows a model theory (the Lee model [166]) which has a renormalized perturbation theory but where the only exactly defined renormalized theory is a free field theory.

(One can obtain an interacting theory for the Lee model but only as the limit of non-Hermitian cutoff interactions. The theory obtained in this limit is the sum of the renormalized perturbation theory. It has undesirable ghost states. The cutoff Hamiltonians of the theories discussed here are all Hermitian (see section 10).)

The idea of fixed points and domains of fixed points is part of the old Gell-Mann–Low renormalization group theory: see [150]. The drawback of the old approach is that it was impossible to locate fixed points or domain boundaries, and in consequence the theory was highly speculative. The calculations described in this paper (and particularly the proposed generalizations of them) depend heavily on the modern approach.

The ϕ^4 theory will be cutoff by introducing a lattice as discussed in section 10. The lattice models of section 10 will be studied for all values of the ϕ^4 coupling constant, i.e. for all values of the s^4 coefficient u_0 in the statistical mechanical analogue. The entire range $0 \leqslant u_0 \leqslant \infty$ is found to lie in the free field domain. The high temperature expansion used here allows one to treat all values of u_0 small or large on an equal footing. The calculations are rough but probably reliable. They have been partly confirmed by calculations with the approximate recursion formula of section 6 (see the end of this section). Further studies will involve adding s^6, s^8, etc. terms to the model; with this added flexibility there is still hope of finding a domain boundary.

Jasnow and Wortis [167] have tried to show the existence of a sharp domain boundary (in another problem) by showing that the critical exponents jump discontinuously across the boundary. However, the high temperature expansion technique they used yields in practical approximations only continuous variable exponents; the existence of the discontinuity is to be inferred. In this section the renormalization group theory will be used to define a function $\psi(u_0)$ which vanishes at the boundary (if there is a boundary). Thus, there is no need to search for discontinuities.

First, the renormalization group analysis will be described. To start with one must understand how trajectories in the free field domain approach the Gaussian fixed point. To illustrate this a section of the critical surface is shown in fig. 13.1. The coordinates are the $s^4(\phi^4)$ coupling

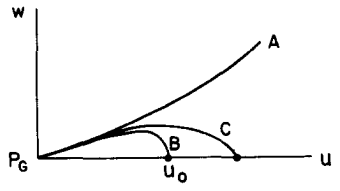

Fig. 13.1. Topology on the critical surface of ϕ^4 field theory. The u axis gives the strength of the bare $u_0\phi^4$ interaction. The w axis is an irrelevant variable. P_G is the Gaussian fixed point. Trajectories B and C are two typical trajectories discussed in the text.

constant u and an irrelevant variable w. The point P_G is the Gaussian fixed point. The initial interactions have no w term and different values of u_0; trajectories for different values of u_0 look like B and C in fig. 13.1. Near P_G they are found to be close to a unique trajectory A. This will be explained later.

If the free field domain has a boundary, then the simplest topological configuration is shown in fig. 13.2: there is a twice unstable fixed point P_B and the boundary of the free field domain is itself the domain of P_B. The trajectory A can now be defined as the trajectory A connecting P_B to P_G. If the boundary intersects the canonical surface (the u axis in this case), then the trajectories from the canonical surface look like D in fig. 13.2 as they approach the boundary.

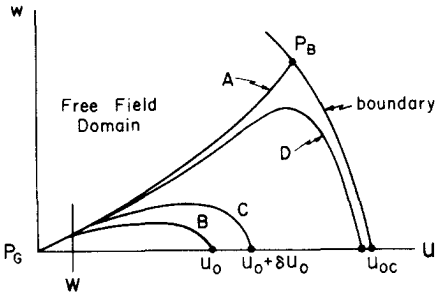

Fig. 13.2. The simplest topology in the presence of a second fixed point P_B. The vertical line W is the gate discussed in the text. u_{0c} is the smallest value of u such that a trajectory does not go to P_G.

The crucial feature of trajectories like D is that they spend a lot of time near the fixed point P_B and therefore do not pass the "gate" W (see fig. 13.2) until very large t. For any initial value u_0 of u, define $t_W(u_0)$ as the time at which the trajectory starting at u_0 passes the gate W. The gate will be placed very close to the Gaussian fixed point, for convenience.

It is easily seen that regardless of the topology, $t_W(u_0) \to \infty$ if u_0 approaches the boundary of the free field domain. The reason is the following. Let u_0 increase. When u_0 reaches the boundary the trajectory leaving u_0 changes discontinuously: instead of going to P_G, it stays on or beyond the boundary. But unless the differential equation defining the trajectories is pathological, only the $t \to \infty$ limit of a trajectory can change discontinuously. Hence, the time to pass the gate W goes to ∞ as u_0 approaches the boundary. This fact will be used to determine whether a boundary exists.

Now define,

$$\psi(u_0) = [\mathrm{d}t_W(u_0)/\mathrm{d}u_0]^{-1}. \tag{13.1}$$

For $t_W(u_0)$ to be infinite at $u_0 = u_{0c}$ while finite for smaller values of u_0, the derivative $\mathrm{d}t_W(u_0)/\mathrm{d}u_0$ must be infinite at $u_0 = u_{0c}$, which means $\psi(u_{0c})$ must be zero. (The function $\psi(u_0)$ is thus analogous to the original ψ function of Gell-Mann and Low in that the existence of a zero of $\psi(u_0)$ at some non-zero value of u_0 makes possible a non-trivial renormalized theory: see [150].) Special attention is required to discuss the point $u_0 = \infty$ in general. However, in this case it will become obvious that $u_0 = \infty$ is inside the free field domain.

It will be seen below that $\psi(u_0)$ is never zero and that at $u_0 = \infty$, ψ is negative. Therefore, $t_W(u_0)$ does not go to ∞ for finite u_0 and is decreasing for large u_0. Therefore, the entire u_0 axis is contained within the free field domain.

First it will be shown that there is a unique trajectory A which is approached for large t by all trajectories originating close to P_G. "Close to P_G" means u and w are small and the perturbation theoretic form of the renormalization group equations is adequate. From section 4, with $d = 4$, the iterative equation for u_l (eq. (4.27), neglecting r_l) is

$$u_{l+1} = u_l - 9cu_l^2. \tag{13.2}$$

Rewritten as a differential equation, this reads,

$$\mathrm{d}u_t/\mathrm{d}t = -9cu_t^2, \tag{13.3}$$

whose solution is,

$$u_t = \frac{1}{9(t - t_0)c} \tag{13.4}$$

where t_0 is a constant. Thus $u_t \to 0$ for $t \to \infty$ like $1/t$. In contrast, a typical equation for an irrelevant variable might be, in lowest nontrivial order,

$$\frac{\mathrm{d}w_t}{\mathrm{d}t} = -2w_t + u_t^2. \tag{13.5}$$

With u_t already known, the solution of this equation is easily obtained:

$$w_t = w_0 \exp(-2t) + \int_0^t \exp\{-2(t - t_1)\}\, u_{t_1}^2 \, \mathrm{d}t_1. \tag{13.6}$$

The only parameter that affects the location of the trajectory is w_0; changing t_0 only changes the origin of t along a trajectory. For large t the w_0 term is completely negligible compared to the second term; hence, except for terms of order $\exp(-2t)$, all trajectories near P_G (on the critical surface) are the same.

The next step is to discuss trajectories which originate very close to the critical surface but not on it. From this analysis a formula for $\psi(u_0)$ will be obtained which can be calculated from a high temperature expansion.

If initially a trajectory is extremely close to the critical surface, then only for very large t will it depart appreciably from the critical surface. Therefore, the trajectory will first come very close to the trajectory A and very close to P_G, but before reaching P_G it leaves the critical surface. The relevant variables near P_G are r and u; possible trajectories leaving the critical surface are shown in fig. 13.3 (B and C). (The parameter r was introduced in section 3. Its initial value r_0 is related to the bare mass μ_0 of the ϕ^4 theory expressed in units of the cutoff: see sections 3 and 10.)

Let the correlation length of the initial interaction be ξ_0. One can vary ξ_0 independently of u_0; the value of ξ_0 is varied by varying $r_0 - r_{0c}(u_0)$ where $r_{0c}(u_0)$ is the value of r_0 on the critical surface.

The function $t_W(u_0)$ can be generalized to a function $t_W(u_0, \xi_0)$ by extending the gate W off the critical surface, as shown in fig. 13.3.

Now a function $\psi(u_0, \xi_0)$ will be defined which is equal to $\psi(u_0)$ for $\xi_0 \to \infty$. If ξ_0 is very large $t_W(u_0 + \delta u_0) - t_W(u_0)$ is only weakly dependent on ξ_0 so that it can be replaced by $t_W(u_0 + \delta u_0, \xi_0 + \delta\xi_0) - t_W(u_0, \xi_0)$, and $\delta\xi_0$ can be chosen to make the calculation simple. Consider the trajectories corresponding to (u_0, ξ_0) and $(u_0 + \delta u_0, \xi_0 + \delta\xi_0)$ – called B and C in fig. 13.3.

The trajectories B and C as shown in fig. 13.3 are distinct beyond the gate. It is convenient to make them coincide; to do this one adjusts $\delta\xi_0$ until B and C meet where they cross the line

$\xi = 1$. Then the trajectories will coincide in the entire region past the gate W. The reason is that specifying $\mathcal{H}_t$ at one point determines the entire trajectory $\mathcal{H}_t$. Actually one cannot make the trajectories coincide *exactly* on the surface $\xi = 1$, because the values of irrelevant variables like w will be slightly different for the two trajectories. In the region before the gate W where u is not small, the irrelevant variables are important and B and C differ appreciably. This was already illustrated in fig. 13.1.

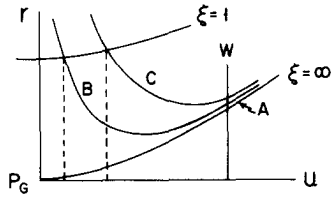

Fig. 13.3. The behavior of renormalization group trajectories (B and C) near P_G and slightly off the critical surface.

The time difference $t_W(u_0 + \delta u_0, \xi_0 + \delta\xi_0) - t_W(u_0, \xi_0)$ is unchanged by moving along the common trajectory from the gate W to the line $\xi = 1$. Thus the time difference can also be written $t_1(u_0 + \xi u_0, \xi_0 + \delta\xi_0) - t_1(u_0, \xi_0)$ where $t_1(u_0, \xi_0)$ is the time the trajectory crosses the surface with $\xi = 1$. Now the time $t_1(u_0, \xi_0)$ is known in terms of ξ_0, since to have $\xi_t = e^{-t}\xi_0 = 1$, one must have

$$t_1(u_0, \xi_0) = \ln \xi_0. \tag{13.7}$$

Therefore, the time difference is,

$$t_W(u_0 + \delta u_0) - t_W(u_0) = \ln(\xi_0 + \delta\xi_0) - \ln \xi_0. \tag{13.8}$$

Let the value of u when $t = t_1(u_0, \xi_0)$ be $u_1(u_0, \xi_0)$. The two trajectories B and C will coincide on the surface $\xi = 1$ if u_1 is the same for both trajectories. Hence $\delta\xi_0$ is to be chosen to keep u_1 constant. Therefore, one defines $\psi(u_0, \xi_0)$ as the derivative,

$$\psi(u_0, \xi_0) = [\mathrm{d} \ln \xi_0 / \mathrm{d}u_0|_{u_1}]^{-1} \tag{13.9}$$

where ξ_0 changes with u_0 such that $u_1(u_0, \xi_0)$ is held fixed. One has $\psi(u_0, \xi_0) \to \psi(u_0)$ for $\xi_0 \to \infty$,

$$\psi(u_0) = \lim_{\xi_0 \to \infty} [\mathrm{d} \ln \xi_0 / \mathrm{d}u_0|_{u_1}]^{-1}. \tag{13.10}$$

The fact that u_0 must vary as ξ_0 varies in order to keep u_1 fixed is an example of coupling constant renormalization. In this example there is a one parameter family of trajectories passing the surface $\xi = 1$ and labelled by the constant u_1; u_1 is closely related to the renormalized coupling constant (see below). As one varies the initial correlation length ξ_0 of a trajectory (which is analogous to varying Λ_0 in section 12) one must also vary u_0 in order that u_1 stay fixed; see section 12.

The next step is to set up a high temperature expansion for $\psi(u_0, \xi_0)$. The high temperature expansions are defined for statistical mechanical models on a lattice. It was explained in section 10 that these lattice models define quantum field theories. The statistical mechanical language will be

sed here. The initial interaction on the lattice is the same one discussed in section 10,

$$\mathcal{H}_0 = K \sum_n \sum_{\hat{i}} s_n s_{n+\hat{i}} - \frac{b}{2} \sum_n s_n^2 - u_0 \sum_n s_n^4 \tag{13.11}$$

/ith K, b, and u_0 being free parameters. The function $\psi(u_0, \xi_0)$ will be obtained from the spin-pin correlation function in the presence of an external field h:

$$\Gamma_n(h, K, u_0) = \langle s_n s_0 \exp(-\mathcal{H}_0) \exp(h\Sigma_n s_n)\rangle / \langle \exp(-\mathcal{H}_0) \exp(h\Sigma_n s_n)\rangle \tag{13.12}$$

vhere

$$\langle\ \rangle \equiv \prod_n \int_{-\infty}^{\infty} ds_n.$$

The high temperature expansion is an expansion of $\Gamma_n(h, K, u_0)$ in powers of K. This expansion ias been discussed extensively (see, e.g., [168]) and will not be reviewed here. The method of 'asnow and Wortis [167] was used to calculate the expansion of $\Gamma_n(h, K, u_0)$ to ninth order in K. The calculation was done (by K.W.) on a computer. There exist programs [169, 170] which :alculate to 12th order in K. When one expands in powers of K, each term involves integrals of :he form,

$$\prod_m \int ds_m P[s_m] \exp\left\{-\frac{b}{2} \sum_n s_n^2 - u_0 \sum_n s_n^4 + h \sum_n s_n\right\} \tag{13.13}$$

vhere $P[s_n]$ is a product of spins. This integral is the product of separate integrals, one for each :attice site; an integral at a lattice site has the form

$$\int_{-\infty}^{\infty} ds\, s^l \exp\{-\tfrac{1}{2} b s^2 - u_0 s^4 + hs\}. \tag{13.14}$$

These integrals were calculated numerically to high precision by the computer.

One is free to make a scale change $s_n \to \zeta s_n$ in the integrals; the only effect is to make a scale change in $\Gamma_n(\Gamma_n \to \zeta^2 \Gamma_n)$ and this will not affect the calculation of $\psi(u_0, \xi_0)$ defined below. Because of this freedom it is unnecessary to calculate for all values of K, b, and u_0. The normalization condition

$$\tfrac{1}{2} b + u_0 = 1 \tag{13.15}$$

was used in the numerical calculation. The spin-spin correlation function was calculated for the values of u_0 shown in table 13.1. It was calculated as an expansion in both K and h.

Table 13.1

u_0	K_c	u_0	K_c
0	0.25	5	0.4218
0.25	0.329	7.5	0.4000
0.5	0.3637	10	0.3794
1	0.4004	20	0.3362
2	0.4276	40	0.3152
3	0.4322	∞	0.3000

The function $\psi(u_0, \xi_0)$ must now be expressed in terms of the spin-spin correlation function. First, change variables from ξ_0 to K. The correlation length ξ_0 is a function of K (for fixed u_0) and goes to ∞ when $K \to K_c(u_0)$ where $K_c(u_0)$ is the critical value of K. So ψ will be considered as a function of u_0 and K ($\psi(u_0, K)$), and the original function $\psi(u_0)$ is,

$$\psi(u_0) = \psi(u_0, K_c(u_0)). \tag{13.16}$$

The definition of ξ_0 to be used here is the "effective range of correlation" introduced in section 3 (eq. (3.30)),

$$\xi_0^2 = -\mathrm{d} \ln \Gamma_q(0, K, u_0)/\mathrm{d}q^2|_{q=0} \tag{13.17}$$

where,

$$\Gamma_q(h, K, u_0) = \sum_n \exp(\mathrm{i}q \cdot n)\, \Gamma_n(h, K, u_0). \tag{13.18}$$

To calculate ψ from eq. (13.10) one needs to know how to keep $u_1(u_0, \xi_0)$ fixed. To do this one needs a quantity which is calculable in terms of $\Gamma_q(h, K, u_0)$ but which depends only on u_1.

Consider the effective interaction $\mathcal{H}_{t_1}$ where a renormalization group trajectory intersects the surface $\xi = 1$. This effective interaction is fixed once $u_1(u_0, \xi_0)$ is determined. Thus, any quantity which depends only on $\mathcal{H}_{t_1}$ will in fact depend only on u_1. The formula for Γ_q for $h = 0$ in terms of $\mathcal{H}_{t_1}$ is therefore

$$\Gamma_q(0, K, u_0) = \exp(-dt_1)\, \zeta^2(u_0, K, t_1) F_2(q \exp(t_1), u_1) \tag{13.19}$$

where,

$$F_2(q, u_1) = Z_1^{-1} \langle \sigma_q \sigma_{q_1} \exp(\mathcal{H}_{t_1}) \rangle / \delta^d(q + q_1), \tag{13.20}$$

Z_1 is the partition function for $\mathcal{H}_{t_1}$, and $\zeta(u_0, K, t_1)$ is the spin renormalization factor. (See eqs. (7.15) and (7.25).) Since $\exp(t_1)$ is ξ_0, one has ($d = 4$ in this section),

$$\Gamma_q(0, K, u_0) = \xi_0^{-4} \zeta^2(u_0, K, t_1) F_2(\xi_0 q, u_1). \tag{13.21}$$

The derivative $\partial^2 \Gamma_q(h, K, u_0)/\partial h^2|_{h=0}$ is a four spin correlation function (with disconnected graphs removed); it is easily seen to have the form

$$\frac{\partial^2 \Gamma_q}{\partial h^2}(0, K, u_0) = \xi_0^{-4} \zeta^4(u_0, K, t_1) F_4(\xi_0 q, u_1). \tag{13.22}$$

(Eqs. (13.21) and (13.22) are true only for large ξ_0: if u_0 is of order 1 and ξ_0 is not large then u_1 is also of order 1. Then $\mathcal{H}_{t_1}$ depends on irrelevant variables as well as u_1; therefore F_2 and F_4 depend on other variables besides $\xi_0 q$ and u_1. This is of no importance here since the equations are needed only for large ξ_0.)

From eqs. (13.21) and (13.22) it follows that the quantity

$$u_R = \xi_0^{-4} \frac{\partial^2 \Gamma_0}{\partial h^2}(0, K, u_0)/[\Gamma_0(0, K, u_0)]^{-2} \tag{13.23}$$

depends only on u_1, namely

$$u_R = F_4(0, u_1)/[F_2(0, u_1)]^2. \tag{13.24}$$

(This u_R is not the same as the u_R of section 9, but they have a similar role – both act as renormalized coupling constants.) It is easy to verify that for small u_1, $F_4(0, u_1)$ is linear in u_1 while $F_2(0, u_1)$ is a constant. Hence, u_R is linear in u_1 to a first approximation. Therefore, holding u_R fixed will hold u_1 fixed.

The formula for $\psi(u_0, K)$ is now,

$$\psi(u_0, K) = [\partial \ln \xi_0/\partial u_0|_{u_R}]^{-1}. \tag{13.25}$$

Since the high temperature expansion gives $u_R(K, u_0)$ and $\xi_0(K, u_0)$, this formula requires a change of variable from K to u_R. The derivative can be expressed in terms of derivatives with u_0 and K as independent variables: one obtains,

$$\psi(u_0, K) = \frac{\partial u_R}{\partial K}\left\{\frac{\partial \ln \xi_0}{\partial u_0}\frac{\partial u_R}{\partial K} - \frac{\partial \ln \xi_0}{\partial K}\frac{\partial u_R}{\partial u_0}\right\}^{-1}. \tag{13.26}$$

Eq. (13.26) was used to obtain the expansion of $\psi(u_0, K)$ in terms of K. This power series was obtained for each of the values of u_0 shown in table 13.1. (The computer calculation was designed to calculate $\partial\Gamma_n/\partial u_0(h, K, u_0)$ as well as $\Gamma_n(h, K, u_0)$.)

In order to compute $\psi(u_0, K_c)$ one must know the critical value $K_c(u_0)$ of K. A standard procedure [168] for determining K_c was used. Consider $\Gamma_q(0, K, u_0)$ for $q = 0$. This is the susceptibility χ. For given u_0, suppose

$$\chi = \Gamma_0(0, K, u_0) = \sum_{l=0}^{\infty} a_l K^l. \tag{13.27}$$

At the critical point the susceptibility diverges (reflecting the onset of spontaneous magnetization). Thus K_c is the radius of convergence of the series in K, unless there are complex singularities closer to the origin. However, all the known coefficients a_l are positive (for $l \leqslant 9$) which suggests that the leading singularity occurs for positive K. Then K_c is given by,

$$K_c = \lim_{l\to\infty} a_{l-1}/a_l. \tag{13.28}$$

The ratios a_{l-1}/a_l are quoted in table 13.2 for $u_0 = 20$. The extrapolation procedures of ref. [168] were used to obtain precise values for $K_c(u_0)$; the error should be less than one percent. The results are quoted in table 13.1.

Table 13.2

l	a_{l-1}/a_l	b_l	l	a_{l-1}/a_l	b_l
0	–	931.88	5	0.328	860892
1	0.283	4743.8	6	0.331	–2857774
2	0.321	–18402	7	0.331	
3	0.323	70212	8	0.332	
4	0.328	–24969	9	0.332	

The power series for $\psi(u_0, K)$ is not rapidly convergent at $K = K_c(u_0)$ so, following an established statistical mechanical practice [168], the sum of the series has been computed by the Padé approximant method. This series for $\psi(u_0, K)$ starts with a K^2 term, so the Padé method was applied to $[\psi(u_0, K)/K^2]$. The procedure is to calculate a "Padé table" at $K = K_c$. The $[M, N]$

entrant in the table is calculated as follows. For a given u_0, let the series expansion for $\psi(u_0, K)$ be,

$$K^{-2}\psi(u_0, K) = \sum_{l=0}^{\infty} b_l K^l. \tag{13.29}$$

(The constants b_l for $u_0 = 20$ are listed in table 13.2; it turns out that one can calculate b_l only for $l \leqslant 6$, given Γ_q to order K^9.) One takes the series to order $M + N$ and expresses it for small K as a ratio of polynomials of order M and N respectively in K, with an error of order K^{M+N+1}:

$$\sum_{l=0}^{M+N} b_l K^l = \left[\sum_{l=0}^{M} c_l K^l\right] \Big/ \left[\sum_{l=0}^{N} d_l K^l\right] + O(K^{M+N+1}). \tag{13.30}$$

Given the constants b_l, this formula determines the $M + 1$ constants c_l and $N + 1$ constants d_l except for a scale factor. Finally, the entrant $[M, N]$ in table 13.3 is,

$$\left[\sum_{l=0}^{M} c_l K_c^l\right] \Big/ \left[\sum_{l=0}^{N} d_l K_c^l\right]. \tag{13.31}$$

Table 13.3

	$N = 0$	1	2	3	4	5
$M = 0$	−1.073					
1	1.546	−2.891				
2	−0.719	−4.119	−3.624			
3	0.836	−3.983	−3.994	−3.870		
4	−0.536	−4.854	−4.097	−2.949	−4.205	
5	0.607	−4.356	−4.518	−4.792	−4.438	−4.342
6	−0.448	−5.093	−4.632	−4.806	−4.806	−3.321

(These numbers calculated for $K_c = 0.3328$ not 0.3362, by mistake.)

All the entries with $M + N \leqslant 6$ were calculated: the Padé table for $u_0 = 20$ is shown in table 13.3. The numbers shown are the approximants for $1000\, K_c^2/\psi(u_0, K_c)$ for $u_0 = 20$. The standard lore [168] about the Padé table is essentially that: (1) the $[M, 0]$ and $[0, N]$ approximants cannot be trusted, (2) entries near the middle of a row are better than near the outside, and (3) there are generally a few "bad eggs" – numbers which make no sense compared to nearby entries. Naturally, the entries near the bottom ($M + N$ large) are better than those near the top. Applying these rules to table 13.3, one concludes that

$$1000\, K_c^2/\psi(20, K_c) = -5 \tag{13.32}$$

with a large error, say 20 percent. There seems to be little chance that $\psi(20, K_c)$ is 0.

The function $\psi(u_0)$ calculated by these procedures is sketched in fig. 13.4 (plotted against $x = u_0/(1 + u_0)$). The function goes to infinity for $u_0 = 5$ and thereafter is negative and finite. From the equation

$$\frac{dt_W}{du_0}(u_0) = \frac{1}{\psi(u_0)} \tag{13.33}$$

one sees that $t_W(u_0)$ increases as u_0 increases; it reaches a maximum at $u_0 = 5$ and then decreases

for $u_0 > 5$. The calculation would have to have a very large error in order that $t_W(u_0)$ go to ∞ for either a finite or infinite u_0. This is not to say an infinite value of $t_W(u_0)$ is ruled out completely; the Padé approach is risky even when 7 terms in the series in K are known.

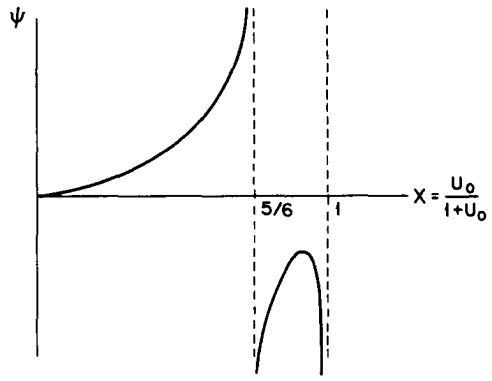

Fig. 13.4. The function ψ whose computation is discussed in the text.

The conclusion of this calculation is that the u_0 axis on the critical surface in the space of initial interactions lies inside the free field domain. It remains to be seen whether the addition of ϕ^6, ϕ^8 terms, etc., to the initial interaction can take one outside the free field domain. Such terms can be studied by the same techniques. One would also like to study more complicated gradient terms ($\phi^2\nabla\phi^2$, etc.) but these terms couple different lattice sites in the lattice models and are harder to study.

The behavior of the ϕ^4 theory in four dimensions was also studied using the approximate recursion formula of section 6. The following was found: For u_0 less than about 0.4, the effective interaction $Q_l(y)$ for large l behaved as $r_l y^2 + u_l y^4$ with u_l small (of order $1/l$). For a large but fixed l, the behavior of u_l versus u_0 was roughly as shown in fig. 13.5: increasing to a maximum value and then decreasing, until at $u_0 \approx 0.4$, u_l goes negative.

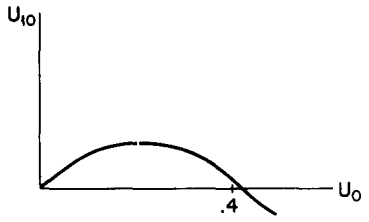

Fig. 13.5. The behavior of u_{10} as a function of u_0 according to the approximate recursion formula.

A negative value of u_l means $\exp\{-Q_l(y)\}$ goes to ∞ for $y \to \infty$, unless there are y^6 terms or higher to stabilize Q_l; this means one has to study the behavior of $Q_l(y)$ for large y, which the author did not do.

The expected size of u_l for large but fixed l, is related to the gate-passing time $t_W(u_0)$. The time $t_W(u_0)$ is the time for which u_t takes on a preassigned value. Since u_t decreases as t increases then if $t_W(u_0)$ increases with u_0, it follows that u_l for fixed l should also increase with u_0; if $t_W(u_0)$

decreases as u_0 increases, then u_l should also decrease. Hence, the rise and fall of u_l in fig. 13.5 is consistent with the high temperature expansion result that $t_W(u_0)$ first increases, then decreases. Because of the many approximations that go into the approximate recursion formula, there is no significance to the fact that the maximum of $t_W(u_0)$ occurs for $u_0 \approx 5$ while the maximum of u_l occurs for $u_0 < 0.4$.

The result from the recursion formula that u_l becomes negative for $u_0 > 0.4$ is not confirmed by the high temperature expansion results. It will be shown below that if the effective coupling u_t becomes negative for finite t for u_0 greater than a "critical" value u_{0t}, then $\psi(u_{0t})$ is zero, and $\psi(u_0)$ is zero, and $\psi(u_0) < 0$ for u_0 slightly less than u_{0t}. In terms of $t_W(u_0)$, this means $t_W(u_0) \to -\infty$ as $u_0 \to u_{0t}$. The high temperature expansion result for ψ (fig. 13.4) shows no evidence for such a zero.

To show that $\psi(u_{0t})$ is zero, consider first the case $u_0 < u_{0t}$. In this case u_t behaves asymptotically for large t as $[(t - t_0)]^{-1}$ (eq. (13.4)) where t_0 depends on u_0. It is easy to see that $t_0(u_0)$ differs only by a constant from $t_W(u_0)$. This is because for $t = t_W(u_0)$, u_t is equal to the value of u at the gate W which is independent of u_0. Therefore, $t_W(u_0) - t_0(u_0)$ is independent of u_0. Hence,

$$\frac{dt_0}{du_0}(u_0) = \frac{1}{\psi(u_0)}.$$

Now, for u_t to decrease as u_0 increases for fixed t, t_0 must be decreasing; for u_t to be zero at $u_0 = u_{0t}$, $t_0(u_{0t})$ must be $-\infty$ and hence $\psi(u_{0t})$ must be zero.

The high temperature expansion method is probably more accurate than the recursion formula. Thus it seems likely that $\psi(u_0)$ is never zero and the entire positive u_0 axis lies in the free field domain.

14. Concluding remarks

In statistical mechanics the renormalization group ideas and the ϵ expansion provide a framework for a more accurate understanding of critical phenomena than Landau theory (or mean field theory). The Landau theory is the first stage in the description of a critical point; the renormalization group and ϵ expansion provide a second stage. Whether more stages are needed is not known at present.

This report has emphasized principles rather than specific applications. Many topics not mentioned here are discussed in the current literature. References were given in section 1.

The relevance of the theory of renormalization in field theory presented here is less clear. It is intended for use in a theory of strong interactions where standard perturbation theoretic renormalization theory could be irrelevant. The exact renormalization group of section 11 provides a basis for the non-Lagrangian formulation of broken scale invariance and current algebra [164] (although not all of the general assumptions of [164] have been derived as yet from the exact equations).

The problems discussed in detail in this review are rather remote from the day-to-day problems of strong interactions. Partly this is because only the high temperature phase in statistical mechanics was discussed. One can expect that in the near future the behavior of super-fluid ^{4}He and the Heisenberg ferromagnet below T_c will be investigated in detail using the new methods. These low temperature phases both involve a broken continuous symmetry. The theory of these phases is likely to show an uncanny similarity to the theory of broken SU(3) x SU(3).

Perhaps the most fundamental problem in strong interactions is to construct constituent models of hadrons. For example one would like to know whether the observed mass spectrum can be obtained as bound states of interacting quarks or other constituents. Is the renormalization group of use here?

At present one does not even know how to study constituent models. Feynman graph expansions assume that all particles are elementary and other approaches are too phenomenological to be relevant here.

The analogy with phase transitions suggests that the place to start is to think in terms of phases. In the ferromagnetic models discussed in this paper there are three phases, one high temperature phase and two low temperature phases. For each phase there is a "phase domain" in the space S of interactions: the interactions in a given phase domain belong to the same phase although their detailed properties vary considerably. For example, the correlation length ξ varies from 0 to ∞ within each phase domain. The analogy to field theory suggests that spaces of field theoretic interactions also divide into phase domains; only some of these domains are easily discussed by Feynman graphs.

The problem is, therefore, to study the possible phases of interacting constituents, to see if any phase resembles the qualitative features of strong interactions. To characterize a phase it is, fortunately, not necessary to study it near the critical point where there is a large correlation length. It is often sufficient to study it in a high or low temperature limit. In field theoretic language this means one can study strongly cutoff field theories. Furthermore, one can use a non-covariant cutoff, in particular a lattice cutoff. It was shown in section 12 how a theory can become covariant near the critical point. (The argument is not guaranteed to work but this is a risk one has to take.) With the simplication of a strong non-covariant cutoff maybe one can develop *honest* methods for solving strongly coupled constituent models and learn what phases are possible. The renormalization group is not involved here because one does not have to study theories with large ξ.

Many questions arise which were not discussed here. Fermions have not been discussed at all. The relation of the new renormalization theory to perturbation theoretic renormalization has been scarcely mentioned. What connections or implications, if any, the renormalization group has for bound states and scattering amplitudes has not been discussed. The idea of field theories in $d < 4$ with anomalous dimensions has not been adequately discussed. Hopefully these questions will be investigated in the near future.

Acknowledgements

The authors are very grateful to Dr. G. Farrar for much assistance in revising the manuscript. We also thank Drs. D. Soper and A. Zee for reading parts of the manuscript.

K. W. has benefited from many discussions with Professor M. E. Fisher, who collaborated in an important part of this work (the ϵ expansion). The author also thanks Professor B. Widom and P. Carruthers for discussions and encouragement. The author enjoyed several conversations with Drs. A. M. Polyakov, A. A. Migdal, and V. N. Gribov, and is especially grateful to Dr. Polyakov for his derivation of the recursion formula (section 6). Professor D. Jasnow's suggestions were very useful in programming the high temperature expansion. Drs. A. Suri and G. Golner contributed to this program as graduate students. Thanks are owed to many others as well.

K. W. wishes to thank Professor C. Kaysen and members of The Institute for Advanced Study and Princeton University for their hospitality during the lectures which led to this report.

Appendix: Simple solutions of the exact renormalization group equations

In this Appendix the free field (Gaussian) fixed point and some of the local "operators" $O_m[x; s]$ associated with the fixed point will be computed for the exact renormalization group equations of section 11. The linearized transformation which determines the operators $O_m[x; s]$ and anomalous dimensions d_m will be derived in general.

Suppose the initial interaction $\mathcal{H}_0$ is pure Gaussian:

$$\mathcal{H}_0 = -\tfrac{1}{2} \int_q \omega(q)\, \sigma_q \sigma_{-q}. \tag{A.1}$$

Then the effective interaction $\mathcal{H}_t$ is also pure Gaussian. The differential equation for $u_2(q, t)$, eq. (11.19), can be written

$$\frac{\mathrm{d}}{\mathrm{d}t} u_2(q\mathrm{e}^t, t) = 2\left[\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q^2 \mathrm{e}^{2t}\right] u_2(q\mathrm{e}^t, t)\,[1 - u_2(q\mathrm{e}^t, t)]. \tag{A.2}$$

The initial condition is

$$u_2(q, 0) = \omega(q). \tag{A.3}$$

The differential equation is separable and easily solved; one obtains

$$\int \mathrm{d}u_2 \left\{\frac{1}{u_2} + \frac{1}{1 - u_2}\right\} = 2\int \left[\frac{\mathrm{d}\rho}{\mathrm{d}t} + 2q^2 \mathrm{e}^{2t}\right] \mathrm{d}t. \tag{A.4}$$

This gives

$$\frac{u_2(q\mathrm{e}^t, t)}{1 - u_2(q\mathrm{e}^t, t)} = \frac{\omega(q)}{1 - \omega(q)} \exp\{2\rho(t) + 2q^2(\mathrm{e}^{2t} - 1)\} \tag{A.5}$$

(assuming that $\rho(0) = 0$). Solving for u_2 gives,

$$u_2(q, t) = \frac{\omega(q\mathrm{e}^{-t})}{\omega(q\mathrm{e}^{-t}) + [1 - \omega(q\mathrm{e}^{-t})] \exp\{-2\rho(t) - 2q^2 + 2q^2 \mathrm{e}^{-2t}\}}. \tag{A.6}$$

To obtain a fixed point one wants to be at a critical point. Since the correlation function $\Gamma(q)$ for the interaction $\mathcal{H}_0$ is $1/\omega(q)$, a critical point occurs if $\omega(0) = 0$. So let

$$\omega(q) = q^2 + \mathrm{O}(q^4) \tag{A.7}$$

for small q. Then to a first approximation for large t and fixed q,

$$u_2(q, t) = \frac{q^2 \mathrm{e}^{-2t}}{q^2 \mathrm{e}^{-2t} + \exp[-2\rho(t) - 2q^2]}. \tag{A.8}$$

For this to approach a fixed function $u_2^*(q)$ as $t \to \infty$ one wants

$$\rho(t) = t + c \tag{A.9}$$

for large t, where c is a constant. Then

$$u_2^*(q) = \frac{q^2}{q^2 + [\mathrm{e}^{-2c}] \exp(-2q^2)}. \tag{A.10}$$

Other choices for $\rho(t)$ lead to t dependence or nonlocal behavior for $u_2(q, t)$ for large t. Nonlocality in position space results from singularities or sudden changes in $u_2(q, t)$ as a function of q. If $\rho(t) \gg t$, for large t, for example, then $u_2(q, t) \approx 1$ except at very small q, where $u_2(0, t) = 0$. So there is an abrupt change in $u_2(q, t)$ at very small q, violating locality. If $\rho(t) \ll t$ for large t, then

$$u_2(q, t) \approx q^2 \exp(2q^2) \exp\{2\rho(t) - 2t\} \tag{A.11}$$

for large t and fixed q and there is no fixed point. So (A.9) is the only acceptable choice. Different choices for c give different values for the term of order q^2 in $u_2(q, t)$. To satisfy the normalization condition

$$u_2^*(q) = q^2 + O(q^4) \tag{A.12}$$

one must choose $c = 0$.

The fixed point is independent of the size of q^4, q^6, etc., terms in $\omega(q)$; the coefficients of q^4, q^6, etc., in $\omega(q)$ are irrelevant variables.

From eq. (11.31) one obtains the renormalization parameter:

$$\zeta(t) = \exp\{(d + 2)t/2\}. \tag{A.13}$$

Comparing with eq. (7.35) one obtains

$$d_s = \tfrac{1}{2}(d - 2) \tag{A.14}$$

which is the canonical dimension for a scalar field. The form (A.13) for $\zeta(t)$ is consistent with the renormalization parameter ζ obtained in section 3, there it was found that $\zeta = 2^{1+d/2}$ for a change in the cutoff by a factor 2. The cutoff e^{-t} changes by a factor 2 when $t \to t + \ln 2$, which in turn increases $\zeta(t)$ by a factor $2^{1+d/2}$.

It is interesting to look directly at the equation for a fixed point function $u_2^*(\boldsymbol{q})$. If $u_2(\boldsymbol{q}, t) = u_2^*(\boldsymbol{q})$ is independent of t, then $d\rho/dt$ must also be a constant, say $d\rho/dt = b$. From eq. (11.19) the equation for $u_2^*(\boldsymbol{q})$ is

$$0 = -\boldsymbol{q} \cdot \nabla u_2^*(\boldsymbol{q}) + 2(b + 2q^2)u_2^*(1 - u_2^*). \tag{A.15}$$

Assume that $u_2^*(\boldsymbol{q})$ depends only on $|\boldsymbol{q}|$. Then

$$q\, du_2^*/dq = 2(b + 2q^2)u_2^*(1 - u_2^*) \tag{A.16}$$

which is an ordinary separable differential equation, with the general solution

$$u_2^*(q)/\{1 - u_2^*(q)\} = \beta^{-1} q^{2b} \exp(2q^2) \tag{A.17}$$

where β is an arbitrary constant; this gives

$$u_2^*(q) = q^{2b}/\{q^{2b} + \beta \exp(-2q^2)\}. \tag{A.18}$$

This reduces to the previous result for $b = 1$ and $\beta = \exp(-2c)$. However, it now appears that there is a fixed point function for any value of b; why should one choose $b = 1$? If b is not an integer, then $u_2^*(\boldsymbol{q})$ is not analytic in the components of the vector $\boldsymbol{q}$ at $\boldsymbol{q} = 0$. In consequence, the interaction in coordinate space is nonlocal. Thus, solutions which are not analytic for $\boldsymbol{q} = 0$ can be rejected. The solutions with integer b are genuine fixed points, but now they form a discrete set. On examination one finds that these fixed points correspond to initial interactions $\omega(q)$ behaving as q^{2b} for small q. (For $b = 0$, the fixed point has $\xi = 0$; for $b > 0$, ξ is ∞.) There is no case where

an initial interaction function $\omega(q)$ analytic at $q = 0$ gives a fixed point nonanalytic in q. Thus, it is consistent to exclude all nonanalytic functions $u_2(q)$ from the space S and hence exclude the cases with b nonintegral.

Next the exact eigenvalue equations for anomalous dimensions will be derived for an arbitrary fixed point $\mathcal{H}^*$. We first consider trajectories of the form

$$\mathcal{H}_t = \mathcal{H}^* + O_m[\sigma] \exp(-d_m t) \tag{A.19}$$

where $O_m[\sigma]$ is a localized functional of σ treated in first order. "Localized" means in principle that $O_m[\sigma]$ depends on the Fourier transform $s(x)$ of σ_q only over a finite region of x. In practice it means that if $O_m[\sigma]$ is expanded in powers of σ, then

$$O_m[\sigma] = \int_{q_1} v_1(q_1)\sigma_{q_1} + \tfrac{1}{2} \int_{q_1}\int_{q_2} v_2(q_1, q_2)\sigma_{q_1}\sigma_{q_2} + \ldots \tag{A.20}$$

where $v_1(q_1)$, $v_2(q_1, q_2)$, etc., are analytic in q_1 and q_2 for real values of the components of q_1 and q_2 (and free of momentum conservation δ-functions, in particular).

Given any localized operator $O_m[\sigma]$ a translated operator $O_m[x; \sigma]$ can be defined as

$$O_m[x; \sigma] = O_m[\sigma'] \tag{A.21}$$

where

$$\sigma'_q = \exp(iq \cdot x)\sigma_q \tag{A.22}$$

(see below for justification) and it will then be shown that

$$\mathcal{H}_t = \mathcal{H}^* + O_m[xe^{-t}; \sigma] \exp(-d_m t) \tag{A.23}$$

is also a trajectory for any x.

Substituting the form (A.19) into the differential equation for $\mathcal{H}_t$, and keeping only terms to first order in O_m, one obtains the fixed point equation for $\mathcal{H}^*$ and an eigenvalue equation for d_m:

$$0 = \int_q \left(\frac{d}{2}\sigma_q + q \cdot \nabla_q \sigma_q\right) \frac{\delta \mathcal{H}^*[\sigma]}{\delta \sigma_q} + \int_q (b + 2q^2) \left\{\frac{\delta \mathcal{H}^*}{\delta \sigma_q}\frac{\delta \mathcal{H}^*}{\delta \sigma_{-q}} + \frac{\delta^2 \mathcal{H}^*}{\delta \sigma_q\, \delta \sigma_{-q}} + \sigma_q \frac{\delta \mathcal{H}^*}{\delta \sigma_q}\right\} \tag{A.24}$$

where $b = d\rho/dt$ must be a constant;

$$-d_m O_m[\sigma] = \int_q \left(\frac{d}{2}\sigma_q + q \cdot \nabla_q \sigma_q\right) \frac{\delta O_m}{\delta \sigma_q}[\sigma] + \int_q (b + 2q^2) \left\{\frac{2\delta \mathcal{H}^*}{\delta \sigma_{-q}}\frac{\delta O_m}{\delta \sigma_q} + \frac{\delta^2 O_m}{\delta \sigma_q\, \delta \sigma_{-q}} + \sigma_q \frac{\delta O_m}{\delta \sigma_q}\right\} \tag{A.25}$$

where we have kept the fixed point value b for $d\rho/dt$ (see section 12). When O_m is expanded in powers of σ (see A.20), one gets the following equations:

$$-d_m v_1(q_1) = A_1 \cdot v_1(q_1) + \int_{q_2} (b + 2q_2^2) v_3(q_1, q_2, -q_2) \tag{A.26}$$

where A_1 is the operator

$$A_1 = -q_1 \cdot \nabla_{q_1} - \tfrac{1}{2} d + (b + 2q_1^2)[1 - 2u_2^*(q_1)] \tag{A.27}$$

(operators A_n will also be used; they are obtained by substituting q_n for q_1 in A_1)

$$-d_m v_2(q_1, q_2) = (A_1 + A_2) \cdot v_2(q_1, q_2) + \int_{q_3} (b + 2q_3^2) v_4(q_1, q_2, q_3, -q_3)$$

$$-d_m v_3(q_1, q_2, q_3) = (A_1 + A_2 + A_3) \cdot v_3(q_1, q_2, q_3)$$
$$+ 2[b + 2(q_1 + q_2 + q_3)^2] v_1(q_1 + q_2 + q_3) u_4^*(q_1, q_2, q_3, -q_1 - q_2 - q_3)$$
$$+ \int_{q_4} (b + 2q_4^2) v_5(q_1, q_2, q_3, q_4, -q_4) \tag{A.28}$$

etc.

Suppose one has solutions $\mathcal{H}^*$ and $O_m[\sigma]$ of eqs. (A.24) and (A.25). Now the translated operators $O_m[x, \sigma]$ will be defined. The definition of a translated operator is that

$$O_m[x; \sigma] = O_m[0; \sigma'] \tag{A.29}$$

if

$$s(y + x) = s'(y) \tag{A.30}$$

where

$$s(y) = \int_q \exp(iq \cdot y)\, \sigma_q, \qquad s'(y) = \int_q \exp(iq \cdot y)\, \sigma'_q. \tag{A.31}$$

Eq. (A.31) implies that $\sigma'_q = \exp(iq \cdot x)\, \sigma_q$. Finally, define $O_m[0; \sigma]$ to be $O_m[\sigma]$.

The operator $O_m[x, \sigma]$ satisfies the equation

$$\nabla_x O_m[x, \sigma] = \int_q iq\, \sigma_q \frac{\delta O_m}{\delta \sigma_q}[x, \sigma]. \tag{A.32}$$

The proof of this is straightforward. From eq. (A.29)

$$\nabla_x O_m[x, \sigma] = \int_q iq\, \sigma'_q \frac{\delta O_m}{\delta \sigma'_q}[0; \sigma'] \tag{A.33a}$$

$$\frac{\delta O_m}{\delta \sigma_q}[x; \sigma] = \exp(iq \cdot x) \frac{\delta O_m}{\delta \sigma'_q}[0; \sigma'] \tag{A.33b}$$

$$\int_q iq\, \sigma_q \frac{\delta O_m}{\delta \sigma_q}[x; \sigma] = \int_q iq\, \sigma'_q \frac{\delta O_m}{\delta \sigma'_q}[0; \sigma']. \tag{A.33c}$$

Comparing eqs. (A.33a) and (A.33c) gives eq. (A.32). Another useful result is that

$$\mathcal{H}^*[\sigma] = \mathcal{H}^*[\sigma']. \tag{A.34}$$

This follows from momentum conservation. From this equation it follows that

$$\frac{\delta \mathcal{H}^*}{\delta \sigma_q}[\sigma] = \exp(iq \cdot x) \frac{\delta \mathcal{H}^*}{\delta \sigma'_q}[\sigma']. \tag{A.35}$$

With these results it will now be shown that eq. (A.23) defines a solution of the renormalization group equation to first order in O_m. If this is true, one must have

$$-d_m O_m[x\mathrm{e}^{-t};\sigma] - x\mathrm{e}^{-t}\cdot\nabla_y O_m[y;\sigma]|_{y=x\exp(-t)} = \int_q \left(\frac{d}{2}\sigma_q + q\cdot\nabla_q\sigma_q\right)\frac{\delta O_m}{\delta\sigma_q}[x\mathrm{e}^{-t};\sigma] + \int_q (b+2q^2)\left\{\frac{2\delta\mathcal{H}^*}{\delta\sigma_{-q}}\frac{\delta O_m}{\delta\sigma_{+q}} + \frac{\delta^2 O_m}{\delta\sigma_q\,\delta\sigma_{-q}} + \sigma_q\frac{\delta O_m}{\delta\sigma_q}\right\}. \tag{A.36}$$

This equation can be verified using eqs. (A.25), (A.32) and (A.35). Only the cancellation of the $\nabla_y O_m[y;\sigma]$ term will be demonstrated explicitly; one has

$$-x\mathrm{e}^{-t}\nabla_y O_m[y;\sigma]|_{y=x\exp(-t)} = -\int_q iq\cdot x\mathrm{e}^{-t}\sigma'_q\frac{\delta O_m}{\delta\sigma'_q}[0;\sigma']. \tag{A.37}$$

One also has (with $y = x\mathrm{e}^{-t}$)

$$\int_q q\cdot\nabla_q\sigma_q\frac{\delta O_m}{\delta\sigma_q}[y;\sigma] = \int_q q\cdot(\nabla_q\sigma_q)\exp(iq\cdot y)\frac{\delta O_m}{\delta\sigma'_q}[0;\sigma'] = \int_q q\cdot(\nabla_q\sigma'_q)\frac{\delta O_m}{\delta\sigma'_q}[0;\sigma'] - \int_q q\cdot iy\,\sigma'_q\frac{\delta O_m}{\delta\sigma'_q}[0;\sigma']. \tag{A.38}$$

The $\nabla_y O_m$ term is cancelled by the second term ($\int_q q\cdot iy\ldots$).

As an example of solutions of the eigenvalue equation for d_m, consider the Gaussian fixed point ($u_2^*(q)$ given by eq. (A.18) with $c = 0$; $u_4^* = u_6^* = \ldots = 0$; and $b = 1$). It is convenient to define an auxiliary function $\psi(q)$ satisfying

$$q\cdot\nabla\psi(q) = \{-1 + (1+2q^2)[1-2u_2^*(q)]\}\,\psi(q). \tag{A.39}$$

A solution of this equation is,

$$\psi(q) = \exp(-q^2)/\{q^2 + \exp(-2q^2)\}. \tag{A.40}$$

Consider now the equations for v_1, v_2, etc. One can easily see that there are solutions with only v_1 or v_2 non-zero; but, if v_3 or v_4 are non-zero, then v_1 or v_2 are also non-zero. The simplest case is if only $v_1(q)$ is non-zero. The equation for $v_1(q)$ then can be written,

$$\{-d_m + q\cdot\nabla_q + \tfrac{1}{2}(d-2)\}\,[v_1(q)/\psi(q)] = 0. \tag{A.41}$$

This equation has the general solution,

$$v_1(q)/\psi(q) = |q|^p f(\hat{q}) \tag{A.42}$$

where p is

$$p = d_m - \tfrac{1}{2}(d-2) \tag{A.43}$$

and $f(\hat{q})$ is an arbitrary function of the unit vector $\hat{q}$ ($\hat{q} = q/|q|$).

One now demands that $v_1(q)$ be analytic at $q = 0$. This condition means that p must be a positive integer or zero (if p is odd then $f(\hat{q})$ must be odd in $\hat{q}$, e.g., $f(\hat{q}) = \hat{q}_i$ where $\hat{q}_i$ is an arbitrary component of $\hat{q}$). Thus, the possible eigenvalues d_m are

$$d_m = \tfrac{1}{2}(d-2) + p \qquad (p \geqslant 0) \tag{A.44}$$

and the first few eigenfunctions are

$$\begin{aligned} d_m &= \tfrac{1}{2}(d-2): && v_{1m}(q) = \psi(q) \\ d_m &= \tfrac{1}{2}(d-2)+1: && v_{1m}(q) = q_i\psi(q) \quad \text{(any } i\text{)} \\ d_m &= \tfrac{1}{2}(d-2)+2: && v_{1m}(q) = q_iq_j\psi(q) \quad \text{(any } i \text{ or } j\text{)} \end{aligned} \tag{A.45}$$

etc.

Construction of solutions with v_2, v_3, or v_4 nonzero is now straightforward (one continues to use the function $\psi(q)$ to simplify the equations). The complete set of eigenvalues d_m, when v_l, v_{l-2}, etc., are nonzero is found to be $d_m = l(\frac{1}{2}(d-2)) + p$ with p an integer (positive or 0). This is in agreement with the known canonical dimensions for a scalar field theory in dimension d.

A word of warning: some of the local operators $O_m[x;\sigma]$ are equivalent to zero, in the following sense. All the n-spin correlation functions $\langle s(x_1)\ldots s(x_n)O_m[x;\sigma]\exp(\mathcal{H}^*)\rangle$ vanish unless x is equal to one of the x_i. An example is the operator $\int_q q^2\psi(q)\sigma_q$ which in field theoretic language turns out to be the operator $\nabla^2\phi$ which vanishes (because a theory described by the Gaussian fixed point is a zero mass free field theory).

References

[1] J. Kondo, Progr. Theor. Phys. (Kyoto) *32* (1964) 37.
[2] P.W. Anderson and G. Yuval, Phys. Rev. *B1* (1970) 1522.
[3] P.W. Anderson, G. Yuval, and D. Hamann, Phys. Rev. *B1* (1970) 4464; Solid State Commun. *8* (1970) 1033.
[4] P.W. Anderson, J. Phys. *C3* (1970) 2436.
[5] A.A. Abrikosov and A.A. Migdal, J. Low Temp. Phys. *3* (1970) 519.
[6] M. Fowler and A. Zawadowski, Solid State Commun. *9* (1971) 471.
[7] M. Fowler, Phys. Rev. *B6* (1972) 3422.
[8] E.A. Guggenheim, J. Chem. Phys. *13* (1945) 253.
[9] L.P. Kadanoff, in Proc. Intern. School of Physics "Enrico Fermi", Corso LI, ed. M. Green (Academic Press, New York, 1971)
[10] E.N. Lorenz, Tellus *16* (1964) 1.
[11] L.D. Landau, Phys. Zurn. Sowjetunion *11* (1937) 26, 545 (English transl.: D. ter Haar, Men of Physics: L.D. Landau, Vol. II (Pergamon, Oxford, 1969)).
[12] E.C.G. Stueckelberg and A. Petermann, Helv. Phys. Acta *26* (1953) 499.
[13] M. Gell-Mann and F.E. Low, Phys. Rev. *95* (1954) 1300.
[14] N.N. Bogoliubov and D.V. Shirkov, Introduction to the Theory of Quantized Fields (Interscience, New York, 1959) Ch. VIII.
[15] L.P. Kadanoff, Physics *2* (1966) 263.
[16] L.P. Kadanoff et al., Rev. Mod. Phys. *39* (1967) 395;
R.B. Griffiths, Phys. Rev. *158* (1967) 176.
[17] K. Wilson, Phys. Rev. *D2* (1970) 1438.
[18] A.Z. Patashinskii and V.L. Pokrovskii, Zh. Eksp. Teor. Fiz. *46* (1964) 994 (Sov. Phys. JETP *19* (1964) 677).
[19] A.M. Polyakov, Zhetf. Pis. Red. *12* (1970) 538 (JEPT Letters *12* (1970) 381).
[20] A.A. Migdal, Phys. Letters *37B* (1971) 386.
[21] G. Parisi and L. Peliti, Lettere Nuovo Cimento *2* (1971) 627;
M. d'Eramo, L. Peliti, and G. Parisi, Lettere Nuovo Cimento *2* (1971) 878;
G. Parisi and L. Peliti, Phys. Letters *41A* (1972) 331.
[22] G. Mack and I.T. Todorov, Phys. Rev. *D8* (1973) 1764.
[23] K. Symanzik, Lettere Nuovo Cimento *3* (1972) 734.
[24] B. Schroer, Berlin preprint (lecture notes) (1972).
[25] G. Mack and K. Symanzik, Commun. Math. Phys. *27* (1972) 247.
[26] G. Mack, in Strong Interaction Physics, eds. W. Rühl and A. Vancira (Springer Verlag, Berlin, 1973) p. 300.
[27] K. Johnson, R. Willey, and M. Baker, Phys. Rev. *163* (1967) 1699.
[28] M. Baker and K. Johnson, Phys. Rev. *183* (1969) 1292; *D3* (1971) 2516, 2541.

[29] S. Adler, Phys. Rev. *D5* (1972) 3021.
[30] C.G. Callan, Phys. Rev. *D2* (1970) 1541.
[31] K. Symanzik, Commun. Math. Phys. *18* (1970) 227; Springer Tracts in Modern Physics *57* (1971) 222.
[32] N. Christ, B. Hasslacher, and A. Mueller, Phys. Rev. *D6* (1972) 3543.
[33] K.G. Wilson, Phys. Rev. *B4* (1971) 3174, 3184.
[34] M.K. Grover, L.P. Kadanoff, and F.J. Wegner, Phys. Rev. *B6* (1972) 311.
[35] M.K. Grover, Phys. Rev. *B6* (1972) 3546.
[36] G.A. Baker Jr., Phys. Rev. *B5* (1972) 2622.
[37] F.J. Dyson, Commun. Math. Phys. *12* (1969) 91.
[38] G. Golner, Phys. Rev. *B8* (1973) 339.
[39] G. Golner, Phys. Rev. *B8* (1973) 3419.
[40] K. Wilson and M.E. Fisher, Phys. Rev. Letters *28* (1972) 240.
[41] F.J. Wegner, Phys. Rev. *B6* (1972) 1891
[42] P. Pfeuty and M.E. Fisher, Phys. Rev. *B6* (1972) 1889.
[43] E.K. Riedel and F.J. Wegner, Phys. Rev. Lett. *29* (1972) 349; Phys. Rev. *B7* (1973) 248.
[44] J.S. Langer and M. Bar-on, Ann. Phys. (N.Y.) *78* (1973) 421.
[45] K.G. Wilson, Phys. Rev. Letters *28* (1972) 548.
[46] P.G. De Gennes, Phys. Letters *38A* (1972) 339.
[47] B. Nickel, Cornell (Clark Hall) preprint; Oxford preprint.
[48] E. Brezin, D. Wallace, and K. Wilson, Phys. Rev. Letters *29* (1972) 591.
[49] E. Brezin, D. Wallace, and K. Wilson, Phys. Rev. *B7* (1973) 232.
[50] S. Ma, Phys. Rev. Letters *29* (1972) 1311.
[51] M. Suzuki, Phys. Letters *42A* (1972) 5.
[52] M. Suzuki, Progr. Theor. Phys. (Kyoto) *49* (1973) 424, 1106, 1440.
[53] M.E. Fisher, S. Ma, and B.G. Nickel, Phys. Rev. Letters *29* (1972) 917.
[54] A.I. Larkin and D.E. Khmel'nitskii, Zh. Eksp. Teor. Fiz. *56* (1969) 2087 (Soviet Phys. JETP *29* (1969) 1123).
[55] C. Di Castro and G. Jona-Lasinio, Phys. Letters *29A* (1969) 322.
[56] F. De Pasquale, C. Di Castro and G. Jona-Lasinio, in Proc. Intern. School of Physics "Enrico Fermi", Course 51, ed. M. Green (Academic Press, New York, 1971).
[57] C. Di Castro, Riv. d. Nuovo Cimento *1* (1971) 199.
[58] A.A. Migdal, Zh. Eksp. Teor. Fiz. *59* (1970) 1015 (Soviet Phys. JETP *32* (1971) 552).
[59] C. Di Castro, Lettere Nuovo Cimento *5* (1972) 69.
[60] F. Wegner, Phys. Rev. *B5* (1972) 4529.
[61] J. Hubbard, Phys. Letters *39A* (1972) 365; *40A* (1972) 111.
[62] J. Hubbard and P. Schofield, Phys. Letters *40A* (1972) 245.
[63] Th. Niemeijer and J.M.J. Van Leeuwen, Phys. Letters *41A* (1972) 211.
[64] Physics Today, "Search and Discovery" (March 1972).
[65] J. Serene and K. Wilson, Cornell (Newman Lab.) preprints.
[66] V.N. Gribov and A.A. Migdal, Zh. Eksp. Teor. Fiz. *55* (1968) 1498 (Soviet Phys. JETP *28* (1968) 784).
[67] A.A. Migdal, Zh. Eksp. Teor. Fiz. *59* (1970) 1015 (Soviet Phys. JEPT *32* (1971) 552).
[68] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *55* (1968) 1026 (Soviet Phys. JETP *28* (1969) 533).
[69] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *57* (1969) 271 (Soviet Phys. JETP *30* (1970) 151).
[70] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *59* (1970) 542 (Soviet Phys. JETP *32* (1971) 296).
[71] M.A. Moore, Lettere Nuovo Cimento *3* (1972) 275.
[72] A. Suri, thesis (Cornell University, 1969, unpublished).
[73] K. Wilson, Phys. Rev. *D3* (1971) 1818.
[74] K. Wilson, Phys. Rev. *D6* (1972) 419.
[75] G. Golner, Phys. Rev. *D8* (1973) 3393.
[76] M.E. Fisher and D.S. Gaunt, Phys. Rev. *133* (1964) A224.
[77] B. Widom, Mol. Phys. *25* (1973) 657.
[78] G. 't Hooft and M. Veltman, Nucl. Phys. *B44* (1972) 189.
[79] C.G. Bollini and J.J. Giambiagi, Phys. Letters *40B* (1972) 566.
[80] J.F. Ashmore, Lettere Nuovo Cimento *4* (1972) 289.
[81] K.G. Wilson, Phys. Rev. *D7* (1973) 2911.
[82] H.E. Stanley, Phys. Rev. *176* (1968) 718.
[83] T.H. Berlin and M. Kac, Phys. Rev. *86* (1952) 821.
[84] M. Kac and C.J. Thompson, Physica Norvegica *5* (1971) 163.

[85] R. Abe, Progr. Theor. Phys. (Kyoto) *48* (1972) 1414; *49* (1973) 113.
[86] R. Abe and S. Hikami, Phys. Letters *45A* (1973) 11.
[87] S. Hikami, Progr. Theor. Phys. (Kyoto) *49* (1973) 1096.
[88] S. Ma, Phys. Rev. *A7* (1973) 2172.
[89] R.A. Ferrell and D.J. Scalapino, Phys. Rev. Letters *29* (1972) 413; Phys. Letters *41A* (1972) 371.
[90] E. Brezin and D. Wallace, Phys. Rev. *B7* (1973) 1967.
[91] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *59* (1970) 542 (Soviet Phys. JETP *32* (1971) 296).
[92] K. Wilson, Phys. Rev. *179* (1969) 1499.
[93] K. Wilson, in Proc. 1971 Intern. Symp. on Electron and Photon Interactions at High Energies, ed. N. Mistry (Laboratory of Nuclear Studies, Cornell Univ., Ithaca, N.Y., 1972) p. 115.
[94] S. Brush, Rev. Mod. Phys. *39* (1967) 883.
[95] M.E. Fisher, Repts. Prog. Phys. *30* (1967) 615.
[96] P. Heller, Repts. Prog. Phys. *30* (1967) 731.
[97] L.P. Kadanoff et al., Rev. Mod. Phys. *39* (1967) 395.
[98] M.E. Fisher, in Lectures in Theoretical Physics (University of Colorado Press, Boulder, Colo., 1965) Vol. VIIC, p. 1.
[99] V.L. Pokrovskii, Usp. Fiz. Nauk. *94* (1968) 127 (Soviet Phys.-Uzpekhi *11* (1968) 66).
[100] Proc. Intern. School of Physics "Enrico Fermi", Course 51, ed. M. Green (Academic Press, New York, 1971).
[101] H.E. Stanley, Introduction to Phase Transitions and Critical Phenomena (Oxford University Press, New York, 1971).
[102] L.D. Landau, Phys. Zurn. Sowjetunion *11* (1937) 26, 545 (English transl.: D. ter Haar, Men of Physics: L.D. Landau, Vol. II (Pergamon, Oxford, 1969)).
[103] L.P. Kadanoff, Physics *2* (1965) 263.
[104] T.H. Berlin and M. Kac, Phys. Rev. *86* (1952) 821.
[105] J.D. Bjorken and S.D. Drell, Relativistic Quantum Fields (McGraw Hill, New York, 1965) p. 181.
[106] R.P. Feynman and A.R. Hibbs, Quantum Mechanics and Path Integrals (McGraw Hill Inc., New York, 1965).
[107] M.E. Fisher, Repts. Prog. Phys. *30* (1967) 615, esp. p. 631.
[108] J.S. Langer, Ann. Phys. *41* (1967) 108.
[109] V.G. Vaks and A.I. Larkin, Soviet Phys. JETP *22* (1966) 678.
[110] A.I. Larkin and D.E. Khmel'nitskii, Soviet Phys. JETP *29* (1969) 1123.
[111] K.G. Wilson, Phys. Rev. *B4* (1971) 3184.
[112] P.W. Anderson, J. Phys. *C3* (1970) 2436.
[113] K. Wilson and M.E. Fisher, Phys. Rev. Letters *28* (1972) 240.
[114] K. Wilson, Phys. Rev. *B4* (1971) 3174.
[115] L.P. Kadanoff, ref. [100].
[116] A.M. Polyakov (private communication).
[117] G.A. Baker Jr., Phys. Rev. *B5* (1972) 2622.
[118] F.J. Dyson, Commun. Math. Phys. *12* (1969) 91.
[119] G. Golner, Phys. Rev. *B8* (1973) 339.
[120] K.G. Wilson, Phys. Rev. *D6* (1972) 419.
[121] K.G. Wilson, Phys. Rev. *140* (1965) B445.
[122] J.S. Langer and M. Bar-on, Ann. Phys. (N.Y.) *78* (1973) 421.
[123] A.I. Larkin and D.E. Khal'nitshii, JETP 29 (1969) 1123.
[124] Proc. Intern. School of Physics "Enrico Fermi", Course 51, ed. M. Green (Academic Press, N.Y., 1973).
[125] A.Z. Patashinskii, Zh. Eksp. Teor. Fiz. *53* (1967) 1987 (Soviet Phys. JETP *26* (1968) 1126).
[126] K. Wilson, Phys. Rev. Letters *28* (1972) 548.
[127] E. Brézin, J.C. Le Guillou, J. Zinn-Justin and B.G. Nickel, Phys. Letters *44A* (1973) 227.
[128] G. 't Hooft and M. Veltman, Nucl. Phys. B44 (1972) 189.
[129] M.A. Moore, D. Jasnow and M. Wortis, Phys. Rev. Letters *22* (1969) 940.
[130] D. Jasnow and M. Fisher (unpublished).
[131] E. Brézin, J.C. Le Guillou and J. Zinn-Justin, Phys. Rev. *D8* (1973) 2418.
[132] K. Wilson, in Proc. 1971 Intern. Symp. on Electron and Photon Interactions at High Energies, ed. N. Mistry (Laboratory of Nuclear Studies, Cornell University, Ithaca, N.Y., 1972) p. 115.
[133] A.H. Mueller and T.L. Trueman, Phys. Rev. *D4* (1971) 1635;
A.H. Mueller, T.L. Trueman, and K.G. Wilson, Phys. Rev. *D5* (1972) 1026.
[134] A.M. Polyakov, Chernogolovka preprint.
[135] P.M. Fishbane and J.D. Sullivan, Phys. Rev. *D4* (1971) 2516.
[136] V.N. Gribov and L.N. Lipatov, Phys. Letters *37B* (1971) 78, and Leningrad preprint.
[137] V.N. Gribov and A.A. Migdal, Zh. Eksp. Teor. Fiz *55* (1968) 1498 (Soviet Phys. JETP *28* (1969) 784).

[138] A.A. Migdal, Zh. Eksp. Teor. Fiz. *59* (1970) 1015 (Soviet Phys. JETP *32* (1971) 552).
[139] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *55* (1968) 1026 (Soviet Phys. JETP *28* (1969) 533).
[140] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *57* (1969) 271 (Soviet Phys. JETP *30* (1970) 151).
[141] A.M. Polyakov, Zh. Eksp. Teor. Fiz. *59* (1970) 542 (Soviet Phys. JETP *32* (1971) 296).
[142] M.A. Moore, Lettere Nuovo Cimento *2* (1972) 275.
[143] A. Suri, thesis (Cornell University, 1969, unpublished).
[144] T.D. Schulz, D.C. Mattis, and E.H. Lieb, Rev. Mod. Phys. *36* (1964) 856.
[145] R.P. Feynman, Rev. Mod. Phys. *20* (1948) 367.
[146] R.P. Feynman and A.R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, Inc., New York, 1965).
[147] M.E. Fisher and R.J. Burford, Phys. Rev. *156* (1967) 583.
[148] L.I. Schiff, Phys. Rev. *92* (1953) 766.
[149] M. Gell-Mann and F.E. Low, Phys. Rev. *95* (1954) 1300.
[150] K. Wilson, Phys. Rev. *D3* (1971) 1818.
[151] K. Wilson, Phys. Rev. *D2* (1970) 1438.
[152] F. Wegner, Phys. Rev. *B5* (1972) 4529.
[153] M.E. Fisher, Repts. Prog. Phys. *30* (1967) 615, pp. 664–666;
M.E. Fisher, S. Ma, and B.G. Nickel, Phys. Rev. Letters *29* (1972) 917.
[154] E.N. Lorenz, Tellus *16* (1964) 1.
[155] K. Wilson, Phys. Rev. *B4* (1971) 3174.
[156] M.A. Moore, Lettere Nuovo Cimento *2* (1972) 275.
[157] L. Kadanoff, in Varenna Summer School Lectures, Course 51, ed. M.S. Green (Academic Press, N.Y., 1971).
[158] R.B. Griffiths, Phys. Rev. Letters *24* (1970) 715.
[159] E. Riedel and F. Wegner, Phys. Rev. Letters *29* (1972) 349.
[160] K. Wilson and M.E. Fisher, Phys. Rev. Letters *28* (1972) 240.
[161] P. Pfeuty and M.E. Fisher, Phys. Rev. *B6* (1972) 1889.
[162] F. Wegner, Phys. Rev. *B6* (1972) 1891.
[163] L. Kadanoff, Phys. Rev. Letters *23* (1969) 1430.
[164] K. Wilson, Phys. Rev. *179* (1969) 1499.
[165] W. Zimmermann, in Lectures on Elementary Particles and Quantum Field Theory, eds. S. Deser, M. Grisaru, and H. Pendleton (M.I.T. Press, Cambridge, 1970); Ann. Phys. (N.Y.) *77* (1973) 536, 570.
[166] T.D. Lee, Phys. Rev. *95* (1954) 1329.
[167] D. Jasnow and M. Wortis, Phys. Rev. *176* (1968) 739.
[168] M.E. Fisher, Repts. Prog. Phys. *30* (1967) 615.
[169] M.A. Moore, D. Jasnow and M. Wortis, Phys. Rev. Letters *22* (1969) 940.
[170] M. Wortis, D. Jasnow and M.A. Moore, Phys. Rev. *185* (1969) 805.

Recent references

This supplement lists more papers on the renormalization group; most of these appeared after the list of references from section 1 [1–94] was completed This list is surely incomplete; we apologize to authors whose work is omitted.

Ref. [1] is an elementary lecture on the renormalization group ideas. S. Ma [2] reviews the renormalization group for large n (n is the number of components of s). Jegerlehner and Shroer [3] review abstract quantum field theory ideas applied to the critical point. Suzuki [4] discusses expansion techniques.

Refs. [5–7] present different formulations of the exact renormalization group.

Niemeyer and Van Leeuwen [8] have formulated and approximately solved a non-perturbative renormalization group transformation for the two dimensional Ising model. They obtain results accurate to about 1%.

Refs. [9–13] are concerned with the ϵ expansion constructed by field theoretic techniques, such as the Callan–Symanzik equations. Parisi [14] applies the Callan– Symanzik equations

directly in 3 dimensions. Fisher and Aharony [15] compute the spin-spin correlation function using the ϵ expansion. (Lebowitz and Penrose [16] prove the exponential decay of correlations away from T_c.) The ϵ expansion is obtained from skeleton graph methods in [17, 18]. The equation of state is computed in [19, 20] (independently of previous references [48, 49]). Refs. [21–24] are also concerned with the ϵ expansion. Ref. [25] reports a calculation on the approximate recursion formula.

Refs. [26–29] are concerned with the $n \to \infty$ limit and $1/n$ expansion. Doniach [26] studies the especially interesting case of d near 2. Critical point dynamics are considered in [30–33]. Refs. [34–37] provide an extensive study of dipolar interactions. Anisotropic interactions (either crystal field terms or directional anisotropies) are considered in [38–45]. A structural transition is discussed in [46]. Long range interactions are considered in [47–49].

Surface effects within the ϵ expansion are discussed by Lubensky and Rubin [50]. Aharony [51] discusses a compressible lattice. Tricritical phenomena are considered in [52–54] and [44]. The properties of systems with $n = -2$ are reported in [55–56]. The relation of $n = 0$ to self-avoiding walks is discussed in [57,58]. The Gaussian model on a lattice is discussed in [59]. Conformal invariance at the critical point (see prior ref. [19]) is discussed in [60, 61].

The behavior of the Potts model at the critical point is puzzling. See prior ref. [39] and refs. [62–64].

The Kondo problem (in its simplest form) has recently been solved by modern renormalization group techniques: see [65].

Ref. [66] is concerned with wave function renormalization. Parisi [67] obtains a relation between anomalous dimensions of tensor operators (see section 9) and the anomalous dimension of ϕ. The ϵ expansion in field theory is discussed in [68–71]. The $1/n$ expansion is applied to field theory in [71].

The idea of asymptotic freedom in quantum field theory is reported in [72–74] (asymptotic freedom refers to theories with free field behavior at short distances but non-trivial interactions at long distances). Two dimensional electrodynamics is discussed in [75].

The renormalization group is formulated in a parton-like language in [76] where the consequences of anomalous dimensions for tensor operators are explained intuitively.

Extensions of the Callan Symanzik and Gell-Mann–Low approach to calculate finite mass corrections at short distances are discussed in [77, 78].

The ϵ expansion is applied to the Gribov Reggeon calculus in [79,80]. A static model in $2 + \epsilon$ dimensions is discussed in [81].

[1] K.G. Wilson, in AIP Conf. Proc., No. 10: Magnetism and Magnetic Materials–1972, eds. C.D. Graham Jr. and J.J. Rhyne (American Institute of Physics, New York, 1973) p. 843.
[2] S.-k Ma, Rev. Mod. Phys. *45* (1973) 589.
[3] F. Jegerlehner and B. Schroer, Acta Physica Austriaca Suppl., to be published.
[4] M. Suzuki, Progr. Theor. Phys. (Kyoto) *50* (1973) 393.
[5] F. Wegner and A. Houghton, Phys. Rev. *A8* (1973) 401.
[6] M. Green, Phys. Rev. *A8* (1973) 1998.
[7] J. Hubbard, J. Phys. *C6* (1973) 2765;
M.J. Bissett, J. Phys. *C6* (1973) 3061.
[8] Th. Niemeyer and J.M.J. Van Leeuwen, Phys. Rev. Letters *31* (1973) 1411; Physica *71* (1974) 17.
[9] E. Brézin, J.C. Le Guillou, and J. Zinn-Justin, Phys. Rev. *D8* (1973) 434, 2418.
[10] E. Brézin, J.C. Le Guillou, and J. Zinn-Justin, Phys. Rev. *B8* (1973) 5330.
[11] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, and B.G. Nickel, Phys. Letters *44A* (1973) 227.
[12] Y. Yamazaki and M. Suzuki, Progr. Theor. Phys. (Kyoto) *50* (1973) 1097.
[13] B. Shroer, Phys. Rev. *B8* (1973) 4200.
[14] G. Parasi, Rome preprint (INFN #46).

[15] M.E. Fisher and A. Aharony, Phys. Rev. Letters *31* (1973) 1238, 1537 (E).
[16] J.L. Lebowitz and O. Penrose, Phys. Rev. Letters *31* (1973) 749.
[17] T. Tsuneto and E. Abrahams, Phys. Rev. Letters *30* (1973) 217.
[18] M.J. Stephen and E. Abrahams, Phys. Letters *44A* (1973) 85.
[19] G.M. Avdeeva and A.A. Migdal, Zh. ETF Pis. Red. *16* (1972) 253 (JETP Letters *16* (1972) 178).
[20] G.M. Avdeeva, Zh. Eksp. Teor. Fiz. *64* (1973) 741.
[21] D.J. Amit, Phys. Letters *42A* (1972) 299.
[22] D.J. Amit, D.J. Bergman, and Y. Imry, J. Phys. *C6* (1973) 2685.
[23] J. Hubbard, Phys. Letters *45A* (1973) 349.
[24] J. Loodts, J.L. Colon, and R. Brout, Phys. Letters *44A* (1973) 453.
[25] Ph. Choquard, R.R. Sari, and R. Calinan, Comp. Phys. Comm. *7* (1974) 69.
[26] S. Doniach, Phys. Rev. Letters *31* (1973) 1450.
[27] R. Abe and S. Hikami, Progr. Theor. Phys. (Kyoto) *49* (1973) 442, 1851.
[28] R. Abe, Progr. Theor. Phys. (Kyoto) *49* (1973) 1074, 1877.
[29] S. Ma, Phys. Letters *43A* (1973) 475.
[30] B.I. Halperin, P.C. Hohenberg, and S.-k Ma, Phys. Rev. Letters *29* (1972) 1548.
[31] M. Suzuki and G. Igarishi, Progr. Theor. Phys. (Kyoto) *49* (1973) 1070.
[32] M. Suzuki, Phys. Letters *43A* (1973) 245.
[33] M. Suzuki, Progr. Theor. Phys. (Kyoto) *50* (1973) 1767.
[34] M.E. Fisher and A. Aharony, Phys. Rev. Letters *30* (1973) 559.
[35] A. Aharony and M.E. Fisher, Phys. Rev. *B8* (1973) 3323.
[36] A. Aharony, Phys. Rev. *B8* (1973) 3342, 3349, 3358, 3363.
[37] A. Aharony, Phys. Letters *44A* (1973) 313.
[38] D.J. Wallace, J. Phys. *C6* (1973) 1390.
[39] I.J. Ketley and D.J. Wallace, J. Phys. *A6* (1973) 1667.
[40] A. Aharony, Phys. Rev. Letters *31* (1973) 1494.
[41] M. Suzuki, Progr. Theor. Phys. (Kyoto) *49* (1973) 1451.
[42] L. Liu, Phys. Rev. Letters *31* (1973) 459.
[43] M.K. Grover, Phys. Letters *44A* (1973) 253.
[44] T.S. Chang and H.E. Stanley, Phys. Rev. *B8* (1973) 4435.
[45] A. Aharony, Phys. Rev. *B8* (1973) 4270.
[46] R.A. Cowley and A.D. Bruce, J. Phys. *C6* (1973) L191.
[47] G.A. Baker Jr. and G.R. Golner, Phys. Rev. Letters *31* (1973) 22.
[48] M. Suzuki, Y. Yamazaki, and G. Igarishi, Phys. Letters *42A* (1972) 313.
[49] J. Sak, Phys. Rev. *B8* (1973) 281.
[50] T. Lubensky and M.H. Rubin, Phys. Rev. Letters *31* (1973) 1469.
[51] A. Aharony, Phys. Rev. *B8* (1973) 4314.
[52] M.J. Stephen and J.L. McCauley Jr., Phys. Letters *44A* (1973) 89.
[53] E.K. Riedel and F.J. Wegner, Phys. Rev. *B9* (1974) 294.
[54] D.J. Amit and C.T. DeDominicis, Phys. Letters *45A* (1973) 193.
[55] R. Balian and G. Toulouse, Phys. Rev. Letters *30* (1973) 544.
[56] M.E. Fisher, Phys. Rev. Letters *30* (1973) 679.
[57] C. Domb, A.J. Barrett, and M. Lax, J. Phys. *A6* (1973) L82.
[58] R.G. Bowers and A. McKerrell, J. Phys. *C6* (1973) 2721.
[59] A. Weyland, Phys. Letters *44A* (1973) 141.
[60] A.M. Wolsky, M.S. Green, and J.D. Gunton, Phys. Rev. Letters *31* (1973) 1193.
[61] A.M. Wolsky and M.S. Green, Phys. Rev. *A9* (1974) 957.
[62] R.J. Baxter, J. Phys. *C6* (1973) L445.
[63] R.J. Baxter and F.Y. Wu, Phys. Rev. Letters *31* (1973) 1294.
[64] H.P. Griffiths and D.W. Wood, J. Phys. *C6* (1973) 2533.
[65] K. Wilson, in Nobel Symposia–Medicine and Natural Sciences, Vol. 24 (Academic Press, New York, 1974).
[66] B. Jouvet, Nuovo Cimento *18A* (1973) 459.
[67] G. Parisi, Nucl. Phys. *B59* (1973) 641.
[68] P.K. Mitter, Phys. Rev. *D7* (1973) 2927.
[69] S.L. Ginzburg, JETP Letters *17* (1973) 448.
[70] S.S. Shei and T.M. Yan, Phys. Rev. *D8* (1973) 2457.
[71] K. Wilson, Phys. Rev. *D7* (1973) 2911.

[72] H. Politzer, Phys. Rev. Letters *30* (1973) 1346.
[73] G.J. Gross and F. Wilczek, Phys. Rev. Letters *30* (1973) 1343; Phys. Rev. *D8* (1973) 3633.
[74] A. Zee, Phys. Rev. *D7* (1973) 3630.
[75] R.J. Crewther, S.S. Shei, and T.M. Yan, Phys. Rev. *D8* (1973) 1730.
[76] J. Kogut and L. Susskind, Phys. Rev. D (to be published).
[77] S. Weinberg, Phys. Rev. *D8* (1973) 3497.
[78] K. Symanzik, Commun. Math. Phys. *23* (1971) 49.
[79] A.A. Migdal, A.M. Polyakov, and K.A. Ter-Martirosyan, Phys. Letters *48B* (1974) 239.
[80] H.D.I. Abarbanel and J.B. Bronzan, Phys. Letters *48B* (1974) 345.
[81] A.A. Belavin and M.A. Yurishchev, Zh. Eskp. Teor. Fiz. *64* (1973) 407.