REVIEWS OF MODERN PHYSICS VOLUME 45, NUMBER 4 OCTOBER 1973

Introduction to the Renormalization Group

Shang-keng Ma[*†]

Department of Physics and Institute for Pure and Applied Physical Sciences, University of California at San Diego, La Jolla, Califonia 92037

The basic idea of the renormalization group is introduced and illustrative examples are presented. Emphasis is put on the application to the theory of critical phenomena. This article is prepared for pedagogical purposes. It is written at a level that a second-year graduate student in physical sciences can understand. No previous knowledge of critical phenomena or field theory is needed. We make no attempt to survey the field or cover a wide range of subjects. On the contrary, we limit the scope to the most basic aspects. We choose to elaborate at length to make the basic idea clear and the definitions precise, and to go through the examples very carefully. We feel that once these basic aspect are understood, there will be no difficulty in confronting the rapidly expanding literature on this subject.

CONTENTS

I. INTRODUCTION

The basic idea of the renormalization group is easy to understand but the mathematical complication involved tends to cover it up. Early advances of the renormalization group idea appeared in relativistic field theory [Gell-Mann and Low (1954)] and in the theory of critical phenomena [Kadanoff (1966)]. More recently, Wilson (1971–1973) has made important progress in bringing the idea into useful and concrete concepts and successfully applied them to different areas of physics. So far the most successful application has been to the theory of critical phenomena. Existing knowledge of critical phenomena has been very helpful in understanding the renormalization group as well. Work in this area has been expanding very rapidly. However, the name 'renormalization group' together with the mathematical complexity going with it has made the subject appear mysterious to many people. The purpose of this article is not to review what has been done. We attempt to introduce this subject at a very elementary level. We shall elaborate only on the most basic ideas and offer simple examples. A more complete coverage at an advanced level does exist [Wilson and Kogut (1972)]. Complete references on the renormalization group can be found there and will not be given here. At the end of this article, we shall give a short guide to the more recent work on this subject.

The conventional formulation of the renormalization group in relativistic field theory [see Bogoliobov and Shirkov (1959), for example] will not be discussed here. The picture and formulation introduced here is based on that put forth by Kadanoff and Wilson. The basic idea behind the two formulations is the same and justifies the same name, even though the two appear very different. The latter is more general and more easily visualized, in our opinion.

We want to do two things in this article. First, we introduce the basic idea and give a precise definition of the renormalization group. Second, we go through two examples. The basic idea is simple and a precise definition is not difficult although it takes a lot of words. The real difficulty is that the properties of the renormalization group are completely unclear from the definition and no classification scheme or rigorous theorem is available. Our general understanding is rather poor at

present. The only way that the reader can see some qualitative features is to go through examples in which everything is worked out explicitly. We shall first tell what to expect and then present the examples. The first example, the infinitely-many-component system or spherical model, requires only patience. The second, the small ϵ case [ϵ=4-dimensional)], is simpler if one knows graph techniques, which we shall explain also. These are the simplest ones that are available.

Our discussion will be within the framework of classical statistical mechanics centered around the theory of critical phenomena. Before giving the outline, it should be helpful to mention a basic ingredient of the renormalization group and to review some qualitative aspects of the theory of critical phenomena.

A. The Role of a Minimum Length

A basic ingredient of the renormalization group is based on the following observation of the role of a minimum-length, which we denote by Λ^{-1} (Λ is thus the maximum wave number or simply the *cutoff* wave number) in describing a given set of physical phenomena. The role of Λ^{-1} is characterized by the facts:

(a) the length scale of the physical phenomena of interest is much larger than Λ^{-1},
(b) the form of the equations and parameters in the equations describing the physical phenomena are defined with respect to Λ^{-1}, and
(c) these parameters summarize the relevant information concerning motions over a scale smaller than Λ^{-1}.

An example will clarify these vague statements. Consider a dilute gas of uranium atoms. In describing the motion of the electrons, i.e., atomic phenomena, the scale of interest is the atomic size, which is much larger than the nuclear size, i.e., $\Lambda^{-1}\sim$nuclear size$\sim$a few fermis. The Schrödinger equation or the Heisenberg equation for the electrons contain parameters such as the total nuclear charge and moments. These are the relevant parameters for phenomena over a scale larger than the nuclear size. They depend on the motion of the 238 nucleons over a scale less than the nuclear size. Note that most of the information concerning the details of the nucleon motion is not relevant for atomic phenomena. Only the over-all features of charge and moments matter. For this Λ^{-1}, the nucleus is a point. Now let us shift our interest to macroscopic phenomena. Suppose we want to describe sound propagation in this uranium gas. The relevant lengths are much larger than the mean free path of the atoms. Thus, we have $\Lambda^{-1}\sim$a few mean free paths$\sim$microns. The equation of interest is the sound-wave equation and the parameters appearing include the compressibility and the viscosity. These parameters can be calculated by studying the motion of atoms over scales less than a few mean free paths. Therefore, we see that when Λ^{-1} changes from a few fermis to a few microns the equations of motion and the parameters change drastically.

The above observations (a), (b), and (c) thus seem too obvious.

Now we ask the nontrivial question of whether we could change Λ^{-1} continuously from a few fermis to a few microns and determine how the Schrödinger equation or the Heisenberg equation together with the parameters change continuously with Λ^{-1} into hydrodynamic equations with compressibility and viscosity appearing as parameters. This appears to be an utterly difficult question and seems pointless, because Λ^{-1} in this example is only a qualitative concept and new equations are constructed everytime a Λ^{-1} is given. However, there are examples where the above observations can be generalized and Λ^{-1} becomes a *quantitative* concept and can be varied continuously. By studying the change of parameters as functions of Λ^{-1}, we can actually learn something. Let us imagine a model of spins on a crystal lattice. We have some coupling parameters describing interactions between neighboring spins. These parameters are defined for Λ^{-1}=one lattice spacing. If we are only interested in spin fluctuations over distances much longer than a lattice spacing, we can introduce a "spin density," which is some kind of average value of spins over a block of size Λ^{-1}. Λ^{-1} can be chosen as 7 or 50 or 350 or any number of lattice spacings as long as it is much less than the length scale we are interested in. The coupling parameters describing the interaction between spin densities on neighboring blocks will depend heavily on Λ^{-1} and the above observations (a), (b), and (c) will apply. The dependence of coupling parameters on the block size Λ^{-1} was studied qualitatively in Kadanoff's theory of critical phenomena. This spin block picture of Kadanoff serves as a starting point of Wilson's formulation of the renormalization group. In short, the renormalization group is a set of transformations of the coupling parameters under changes of Λ and some other changes. Why such transformations are relevant to physics is, of course, the subject we shall explain and illustrate.

To define such transformations in a precise and useful manner is not easy even though the basic idea is simple. For a given block size Λ^{-1}, we are thinking of all the spins within each block as behaving like one unit. Each unit interacts with neighboring units. While it is easy to define nuclear charges and moments, the "average spin" over a block and the effective interaction parameters for neighboring blocks are not easy to obtain. The formal definition for the renormalization group introduced here will be based on that of Wilson (1971). It is a bit lengthy, but the qualitative features are just those given above, and should be kept in mind as the discussion becomes formal.

How can we learn anything by changing the block size? We simply get a new set of parameters and in general we do not gain anything. However, there are

cases where the change of parameters has a tractable pattern and we can then get somewhere. In applications to critical phenomena, for example, a simple pattern does seem to emerge.

B. Some Qualitative Features of Critical Phenomena and the Scaling Hypothesis

If the reader is already familiar with the general features of critical phenomena and scaling arguments, he can go directly to Sec. IC. For a complete review, see the works of Kadanoff *et al.* (1967) and Fisher (1967).

Let us imagine a sample of an isotropic ferromagnetic material. If the temperature T is below its critical temperature T_c, there is a spontaneous magnetization. Right above T_c, there is not. There are large fluctuations in magnetization for T near T_c. As the temperature T approaches T_c, the magnetic susceptibility and some other measurable quantities diverge. For example, the susceptibility diverges like $(T-T_c)^{-\gamma}$, for $T>T_c$, where γ, one of the *critical exponents*, is observed to be near 1.3 for many materials exhibiting a critical point. There are many other critical exponents describing the divergence of other quantities at the critical point. The theory of critical phenomena has the task of explaining these divergences.

These divergences are believed to be consequences of the large fluctuations of magnetization. Also, the observed universal (i.e., independent of materials) character of these divergences suggests that only the large scale behavior, not the detailed microscopic interactions, is relevant in a correct explanation.

The *scaling hypothesis* plays an important role in the theory of critical phenomena. This hypothesis says that there is a *correlation length* ξ (which may be thought of as measuring the average distance over which the fluctuations of magnetization are correlated) which is the longest and the only relevant length in explaining critical phenomena. Other lengths, such as the interatomic distances, are too short to play a role. The hypothesis says also that ξ, diverging like $|T-T_c|^{-\nu}$, $\nu>0$, accounts for the dominating temperature dependence near T_c of all quantities. In other words, physical quantities depend on $T-T_c$ only through their dependence on ξ. For example, it leads to the following very important consequence. If we increase the unit of length by a factor s, then in the new unit, the system appears *shrunk* by a factor s. The correlation length now becomes ξ/s under this scale change. Since the correlation length is proportional to $|T-T_c|^{-\nu}$, a decrease in correlation length corresponds to an increase in $|T-T_c|$. Therefore, near T_c the temperature dependence of a physical quantity can be deduced from the way it behaves under a change of scale. The simplest example applying this idea is the following. The free energy per unit volume $F(\xi)$ becomes $s^dF(\xi)$ when the volume of the system is shrunk; d is the dimension. Therefore we have $F(\xi/s)=s^dF(\xi)$. Since s is arbitrary, we set $s=\xi$. We then have

$$F(\xi)=s^{-d}F(\xi/s)=\xi^{-d}F(1)\propto|T-T_c|^{\nu d}, \qquad (1.1)$$

since $\xi\propto|T-T_c|^{-\nu}$. Such arguments do not sound too convincing. Later, we shall be more critical. Another important consequence is that in the limit $T\to T_c$, ξ becomes infinite and there is no longer any length parameter. Thus the system would look the same if a change in length scale is made. There are many important consequences of the scaling hypothesis as well as many ambiguities. It is clear that this hypothesis is very powerful but that its origin is not clear. A more fundamental understanding is needed. We shall see how it can be understood in the light of a renormalization group analysis.

C. Outline

The first half of this article (Secs. I–III) is devoted to the explanation of basic ideas of the renormalization group and to its formalism. The other half offers illustrative examples. Our discussion will be centered around the statistical mechanics of an n-component classical field in a d-dimensional space. Usually the dimension of physical interest is $d=3$. The number of components n is 3 if the classical field is to describe the magnetization in a ferromagnet. Near the λ transition of liquid ^{4}He, the quantum amplitude of the ^{4}He atoms is expected to be describable as a classical field of $n=2$. (The amplitude is a complex number, which has two real components, namely the real part and the imaginary part.)

Two examples will be illustrated here, the case of large n (Sec. IV), which is equivalent to the spherical model [see Stanley (1968)], and the case of small $4-d\equiv\epsilon$ [Wilson and Fisher (1972)] (Sec. VI). The unphysical condition of large n or small ϵ is necessary to simplify the mathematics.

In Sec. V, an elementary introduction to graph expansion will be given. The small ϵ case in Sec. VI will be discussed with the help of graphs. A great deal of use of graph expansion has been made in recent works on this subject. That is why we introduce it here for those who have not been exposed to it before. However, we want to emphasize that one of the most valuable features of the renormalization group formalism is that it is free from any perturbation theory. The graph expansion, which is a perturbation expansion, is *not essential* to the renormalization group, in spite of the fact that it is helpful in some cases. All of our discussion on the basic ideas, the formulation, the illustration of the large n case (Secs. II–IV), will be completely free from graphs. If the reader understands the material all the way up to where the graphs start, he has understood the essentials which we want to present.

We should point out that the formulation of the renormalization group is by no means unique. There is a great deal of flexibility which is unexplored so far. Our

emphasis on critical phenomena in this article does not imply that the renormalization group applies only to critical phenomena. There are many areas of physics where the renormalization group can be useful.

In Sec. VII, the basis for perturbation expansions of critical exponents is discussed. Existing expansions are those in powers of ϵ and those in powers of $1/n$.

Some general remarks are made in Sec. VIII, and a brief guide to recent literature on renormalization group is included.

II. THE RENORMALIZATION GROUP DEFINED

A renormalization group can be defined for any large system such as a thermodynamical system or a quantum field. We shall define a renormalization group for a model thermodynamical system analyzed in the framework of classical statistical mechanics. But before we proceed with our definitions, we would like to remind the reader of some truly trivial facts concerning probability distributions.

A. Digression on Trivial Observations

Let $P(y_1, y_2, y_3)$ be the probability distribution function for the random variables $-\infty < y_1, y_2, y_3 < \infty$. To calculate the average value of any function $f(y_1, y_2, y_3)$ of these random variables, for example, $f = y_1 y_2$, we simply do the integral

$$\langle f \rangle_P = \langle y_1 y_2 \rangle_P = \int_{-\infty}^{\infty} dy_1\, dy_2\, dy_3 y_1 y_2 P(y_1, y_2, y_3). \quad (2.1)$$

We notice that for those f which do not depend on y_3, we can obtain an equivalent distribution function $P'(y_1, y_2)$ by integrating out the variable y_3 from $P(y_1, y_2, y_3)$, i.e.,

$$P'(y_1, y_2) \equiv \int_{-\infty}^{\infty} dy_3 P(y_1, y_2, y_3). \quad (2.2)$$

Therefore, let us remember

Fact 1: P', obtained from P by integrating out certain random variables, is equivalent to P provided we are not interested in these integrated variables. Next, we observe that if we obtain a new probability distribution $P'(y_2, y_4, y_6)$ from $P(y_1, y_2, y_3)$ by changing the name of the random variables, we won't get anything new. For example we have

$$P'(y_2, y_4, y_6) \equiv P(y_2, y_4, y_6), \quad (2.3)$$

i.e., replacing 1, 2, 3 in $P(y_1, y_2, y_3)$ by 2, 4, 6. The only thing we must watch out for is that when we calculate averages we must change labels accordingly. For example we have

$$\begin{aligned} \langle y_2 \rangle_P &= \int dy_1\, dy_2\, dy_3 P(y_1, y_2, y_3) y_2, \\ &= \int dy_2\, dy_4\, dy_6 P'(y_2, y_4, y_6) y_4, \\ &= \langle y_4 \rangle_P, \end{aligned} \quad (2.4)$$

i.e., we must calculate the average of y_4 over P' if we want to get the average of y_2 over P. This sounds too trivial, but must be remembered.

Fact 2: P', obtained from P by relabeling the random variables, is equivalent to P provided that when average values are computed we relabel the random variables of interest accordingly. Finally, if α is a positive constant and

$$P'(y_1, y_2, y_3) \equiv \alpha^3 P(\alpha y_1, \alpha y_2, \alpha y_3) \quad (2.4')$$

then P' clearly says nothing new. Any average calculated over P' is easily related to that over P. For example we have

$$\langle y_1 \rangle_P = \alpha \langle y_1 \rangle_{P'}, \qquad \langle y_1^2 \rangle_P = \alpha^2 \langle y_1^2 \rangle_{P'}. \quad (2.5)$$

Therefore, let us remember

Fact 3: P', obtained from P by changing random variables by a constant factor, is equivalent to P provided we multiply the random variables of interest by the same factor when average values are computed.

We list these three trivial observations so that it will be easier for the reader to understand the more complicated, but basically the same, procedures later. A transformation in the renormalization group essentially transforms a given probability distribution to an equivalent one by the above mentioned three steps: integration, relabeling, and multiplication of random variables by a constant.

B. Model and Notation

Imagine a d-dimensional crystal lattice of volume L^d, where L is measured in units of lattice spacing. At each lattice site x, there is an n-component vector "spin" $\phi(x) \equiv [\phi_1(x), \phi_2(x) \cdots \phi_n(x)]$. Let ϕ_k denote the Fourier components of $\phi(x)$:

$$\phi_i(x) = L^{-d/2} \sum_k \phi_{ik} \exp(ik \cdot x), \quad (2.6)$$

where the sum over wave vectors k is taken over the L^d discrete points in the first Brillouin zone. The density of points, $L^d(2\pi)^{-d}$, is very large since L is a very large number. Each ϕ_{ik} is regarded as a random variable. There are nL^d of them. The probability distribution for these random variables is given by

$$P_{\text{micro}} \propto \exp(-H_{\text{micro}}/T), \quad (2.7)$$

where T is the temperature and H_{micro} is the Hamiltonian which is assumed to be a given function of all the random variables. We assume that H_{micro} is invariant under rotation in the n-dimensional spin vector space and under translation in x space.

The *correlation function* $G(k)$ is defined as

$$\begin{aligned} G(k)\,\delta_{ij} &= \int d^d x \langle \phi_i(x) \phi_j(0) \rangle \exp(-ik \cdot x), \\ &= \langle |\, \phi_{ik}\,|^2 \rangle \delta_{ij}, \end{aligned} \quad (2.8)$$

where the average $\langle \cdots \rangle$ is taken over P_{micro} as given by (2.7). If a term

$$-\int d^d x \phi_1(x) H \quad (2.9)$$

is added to the Hamiltonian, i.e., when a "magnetic field" H is turned on in the 1 direction, we can define the susceptibility as

$$\partial\langle\phi_1\rangle/\partial H. \tag{2.10}$$

It is very easy to show that the susceptibility is just $G(0)$. Other quantities of interest will be defined later.

Since the probability distribution is assumed to be invariant under rotation in spin space, we expect $G(k)$ to be independent of i if there is no external field. However, a rotationally invariant probability distribution can still produce average values which are not rotationally invariant. This happens below T_c, where one of the components, say ϕ_1, has a nonzero average even when $H\to 0$. In our discussions, we shall always assume that $H=0$ unless otherwise specified.

C. The Idea of an Effective Hamiltonian

What we are interested in is the behavior of long-wavelength fluctuations, i.e., that of ϕ_k with small k. The Hamiltonian is usually given by nearest-neighbor interactions. Since we expect the characteristics of long wavelength fluctuations to be independent of the microscopic details, we should be able to obtain an effective Hamiltonian with these irrelevant details removed. In other words, this effective Hamiltonian should not involve any ϕ_k with large k. Of course, the effective Hamiltonian must lead to the same results as the original Hamiltonian would have when averages involving ϕ_k's with small k are calculated. How do we find this effective Hamiltonian? It is very easy in principle. Remember the trivial Fact 1 mentioned at the beginning of this section: We may simply integrate out the irrelevant random variables. Thus, P_{micro}, as given by (2.7), is equivalent to, apart from a normalization constant,

$$\prod_{i,k>\Lambda}\int d\phi_{ik}\exp\left(-H_{\text{micro}}/T\right) \equiv \exp\left[-H(\Lambda)/T\right], \tag{2.11}$$

where the multiple integral is taken over all ϕ_{ik}'s with all $i=1,\cdots,n$ and all k larger than Λ. The cutoff Λ is taken to be much smaller than the inverse lattice spacing but still much larger than the small range of k which is of ultimate interest. $H(\Lambda)$ defined by (2.11) is the desired effective Hamiltonian. Note that we set Λ this way to leave the ϕ_k's in the intermediate k range unintegrated. This is because, besides the random variables in the small k range themselves, those in the intermediate k range also play an important part in determining the small k behavior. The effective Hamiltonian $H(\Lambda)$ tells us about the interactions down to a minimum distance Λ^{-1}. The finer details beyond this distance are averaged out. The multiple integrals in (2.11) will not be easy to carry out explicitly. However, we expect that $H(\Lambda)$ in general will look very different. For example, if the microscopic Hamiltonian has only quadratic and quartic terms in ϕ, the multiple integral of (2.11) will generate all powers of ϕ for $H(\Lambda)$. This will become more evident later. The important point to remember is that the cutoff Λ is an inseparable part of the definition of a Hamiltonian. The fluctuations over a distance less than Λ^{-1} play a role in determining the structure of $H(\Lambda)$.

The ultimate task is to derive singular behavior of physical quantities such as the correlation function near the critical point from a generally nonsingular Hamiltonian. Constructing $H(\Lambda)$ does not seem to help in this task. No singularity is expected in $H(\Lambda)$ since we smeared out fluctuation only over very short distances. If we are now to study critical behavior starting from $H(\Lambda)$, then the task would appear to be much worse than before because $H(\Lambda)$ would look far more complicated than the microscopic Hamiltonian. However, we will be able to see the major characteristics of the critical behavior, which are independent of the details of the details of $H(\Lambda)$, by examining how $H(\Lambda)$ would behave under the renormalization group, which is a set of transformations and will be defined shortly.

The quantity Λ^{-1} plays the role of the size of the spin block discussed in the Introduction. Of course, we are not constructing spin blocks explicitly as was said there. We are simply integrating out fluctuations of wavelengths shorter than Λ^{-1}. Mathematically this is a simpler procedure than is constructing block spin variables. The effect is the same.

D. The Parameter Space

We shall now be more general and consider a large class of probability distributions for ϕ_{ik}. We forget about our spin model introduced above and regard the ϕ_{ik}'s just as a set of random variables. But we still want the label k to range over discrete points in a sphere of radius Λ in k space. The density of points is $L^d(2\pi)^{-d}$. Of course, we have $1\le i\le n$, as before.

Any probability distribution for these random variables can be specified by a set of parameters. Let us imagine that each set of parameters is a point in a *parameter space*, so that any probability distribution P is represented by a point μ in this space. Let us consider a simple example of a parameter space. We write

$$P\propto\exp\left(-\mathcal{H}\right)$$
$$\mathcal{H}=\int d^dx\left[a(\nabla\phi)^2+t_0\phi^2+\tfrac{1}{2}u_4\phi^4\right], \tag{2.12}$$

where we use the abbreviation

$$(\nabla\phi)^2\equiv\tfrac{1}{2}\sum_{i=1}^{n}\left[\nabla\phi_i(x)\right]^2,$$
$$\phi^2\equiv\tfrac{1}{2}\sum_{i=1}^{n}\left[\phi_i(x)\right]^2,$$
$$\phi^4\equiv(\phi^2)^2,$$
$$\phi_i(x)\equiv L^{-d/2}\sum_{k<\Lambda}\exp\left(ik\cdot x\right)\phi_{ik}. \tag{2.13}$$

Equation (2.12) is the so-called "Landau–Ginzburg" form for the probability distribution. In terms of ϕ_k, we can write $\mathcal{H}$ as

$$\mathcal{H}=\tfrac{1}{2}\sum_{k,i}|\phi_{ik}|^2(t_0+ak^2)+\tfrac{1}{8}L^{-d}\times\sum_{k,q,p}\sum_{i,j}u_4\phi_{ip+k}\phi_{iq-k}\phi_{jp}\phi_{jq}. \quad (2.14)$$

All wave vectors are restricted to less than Λ. The quantities a, t_0, and u_4 are our coupling parameters and our parameter space is a three-dimensional space of triplets:

$$\mu=(a, t_0, u_4). \quad (2.15)$$

Namely, given a point in this space, there corresponds a probability distribution via (2.12). Note that the factors $\frac{1}{2}$ thrown into (2.12) and (2.13) are purely a matter of convention, which can be changed at any time for whatever reason. Of course, not all points in the parameter space we just defined are allowed. For example, if u_4 is negative, the probability distribution could not be integrated and therefore would be meaningless. It turns out that this 3-dimensional parameter space is large enough for a nontrivial realization of the renormalization group transformations only for small ϵ, i.e., for d very close to 4. Otherwise, this parameter space is not big enough. In other words, under a renormalization group transformation, the transformed $\mathcal{H}$, which will be defined later, will in general have more terms than the Landau–Ginzburg form gives. For example, there may be terms such as $(\phi^2)^3$, $(\phi^2)^4$, etc. We have to enlarge our parameter space. The qualitative features are already in the space defined via Eq. (2.15). All that is needed is a straightforward generalization. We write

$$P\propto\exp(-\mathcal{H}),$$

$$\mathcal{H}=\sum_{m=1}^{\infty}L^{-(m-1)d}\sum_{k_1,k_2\cdots k_{2m-1}}\sum_{i_1,i_2\cdots i_{2m}}\phi_{i_1k_1}\phi_{i_2k_2}\cdots\phi_{i_{2m}k_{2m}}\times u_{2m}+\text{constant}, \quad (2.16)$$

where $k_{2m}=-(k_1+k_2+\cdots+k_{2m-1})$, and u_{2m} is a function of k_1, $k_2\cdots k_{2m-1}$ and of i_1, $i_2\cdots i_{2m}$. Or, in the coordinate representation, we have

$$\mathcal{H}=\sum_{m=1}^{\infty}\sum_{i_1,i_2\cdots i_{2m}}\int d^dx_1\cdots d^dx_{2m}\phi_{i_1}(x_1)\phi_{i_2}(x_2)\cdots\phi_{i_{2m}}(x_{2m})\times v_{2m}(x_1-x_{2m}, x_2-x_{2m}\cdots x_{2m-1}-x_{2m}), \quad (2.17)$$

where v_{2m} are related to u_{2m} via

$$u_{2m}=\int\prod_{l=1}^{2m-1}[d^dy_l\exp(-ik_l\cdot y_l)]\times v_{2m}(y_1, y_2\cdots y_{2m-1}). \quad (2.18)$$

We shall assume that v_{2m} represents short range interactions (i.e., $v_{2m}\to 0$ if one or more of the y's becomes large) so that u_{2m} can be expanded in powers of k.

We now define our parameter space as the space of all possible μ,

$$\mu\equiv(u_2, u_4, u_6, \cdots). \quad (2.19)$$

This is the generalization of (2.15). Note that each of the entries in (2.19) can contain more than one parameter. For example, the parameters a, t_0 of (2.15) are contained in u_2 of (2.19), i.e.,

$$u_2=\tfrac{1}{2}(t_0+ak^2)\delta_{i_1i_2}.$$

Now the parameter space is enormous. Again, not all of the space is of interest. Also, further generalization may still be necessary. Symmetry considerations will limit the allowable range of the u_{2m}'s. An external field will necessitate the inclusion of odd powers of ϕ in $\mathcal{H}$ as well. In any case, further restrictions and adjustments can always be made when necessary. For qualitative discussions, we can simply think in terms of (2.15). For more formal discussions, we must use (2.16)–(2.19).

We do want to emphasize that Λ, the cutoff in k space, is, unless otherwise specified, always fixed for all probability distributions. The coupling parameters are meaningless unless Λ is fixed. Another important point is that L, which tells us how many random variables there are, is not included as a parameter. This is because we are interested in the limit of infinite L. Averages of interest are always L-independent in this limit. In fact we shall write $\mu=\mu'$ as long as $u_{2m}=u_{2m}'$ for all m even if $L\neq L'$.

Finally, to those readers who are too used to statistical mechanical terminology, we want to emphasize that $\mathcal{H}$, defined by (2.16), is not to be thought of as "energy divided by temperature." It is just the logarithm of the probability distribution. As far as our parameter space is concerned, the concepts of energy and temperature are irrelevant. They enter only in (2.11) as inputs in determining a particular probability distribution corresponding to a particular point in the parameter space.

E. The Renormalizaton Group

Consider the following transformation which takes a probability distribution P to another probability distribution P'. We want to represent this transformation as

$$\mu'=R_s\mu, \quad (2.20)$$

which transforms the point μ to μ' in the parameter space. Of course, μ and μ' represent P and P', respectively. This transformation R_s is defined implicitly by

$$P'\propto\exp(-\mathcal{H}')=\Big[\prod_{i,\Lambda/s<k'<\Lambda}\int d\phi_{ik'}\exp(-\mathcal{H})\Big]_{\phi_k\to\alpha_s\phi_{sk}}, \quad (2.21)$$

where sk means s times k. Equation (2.16) defines μ, and μ' is to be extracted from $\mathcal{H}'$ by writing $\mathcal{H}'$ in the form of Eq. (2.16) and identifying the coefficients of products of random variables. Three steps are involved

in (2.21). First, we integrate out those $\phi_{k'}$ with k' between Λ/s and Λ. Second, we relabel the random variables by enlarging the wave vectors by a factor s. Third, we multiply all random variables by a constant factor α_s. The three trivial facts listed at the beginning of this section imply that P' is equivalent to P as far as random variables ϕ_k with $k<\Lambda/s$ are concerned and provided that proper relabeling and multiplying by α_s are done when averages are computed. For example we have

$$\langle|\phi_{ik}|^2\rangle_P=\alpha_s^2\langle|\phi_{isk}|^2\rangle_{P'}. \tag{2.22}$$

If we define $G(k,\mu)\equiv\langle|\phi_{ik}|^2\rangle_P$, Eq. (2.22) says

$$G(k,\mu)=\alpha_s^2G(sk,R_s\mu). \tag{2.23}$$

Note that the number of random variables in P' is smaller by a factor s^{-d} than that in P owing to the multiple integral in (2.21). The change of scale $k\to sk$ makes the density of points in k space smaller by the same factor. These simply mean that the volume of the system described by P' is $L'^d\equiv s^{-d}L^d$, i.e., shrunk by a factor s^{-d}. To identify μ' from $\mathfrak{H}'$ given by (2.21) we must write $\mathfrak{H}'$ in the form of (2.16) with L' replacing L; the density of points in k space is now $L'^d(2\pi)^{-d}$. As was mentioned earlier, L' or L plays no role in calculating quantities of interest and is not included as a parameter. The set of R_s, $1\leq s<\infty$, will be called the "renormalization group." We did not define the inverse of R_s; thus it is not quite a group.

In terms of our spin block picture discussed in the introduction, what (2.21) does is just to increase the size of a spin block from Λ^{-1} to $s\Lambda^{-1}$ and then change the length unit so that the size of a spin block returns to Λ^{-1} again. As a result, the parameters change from μ to $\mu'=R_s\mu$.

So far nothing has been said about the α_s in (2.21). The only role of α_s is in the last substitution in (2.21). If we have two successive transformations R_s and $R_{s'}$, then it is clear from (2.21) that they have the same effect as a single transformation $R_{ss'}$ except that the substitution is $\phi_k\to\alpha_s\alpha_{s'}\phi_{ss'k}$, and not $\phi_k\to\alpha_{ss'}\phi_{ss'k}$. Thus, in order to observe

$$R_sR_{s'}\mu=R_{ss'}\mu \tag{2.24}$$

for any μ, we must demand

$$\alpha_s\alpha_{s'}=\alpha_{ss'}. \tag{2.25}$$

We shall so restrict our choice of α_s. Equation (2.25) is a severe restriction. It requires that

$$\alpha_s=s^y, \tag{2.26}$$

where y is a constant. If we regard the substitution

$$\phi_k\to s^y\phi_{sk} \tag{2.27}$$

in (2.21) as a scale change, then y can be interpreted as the dimension of ϕ_k in units of length. The dimension of ϕ_k can be defined by the microscopic Hamiltonian. However the dimension so defined is not useful. Instead, we shall determine y with respect to a *fixed point.*

A fixed point μ^* in the parameter space is that satisfying

$$R_s\mu^*=\mu^*. \tag{2.28}$$

It will play a major role in later discussions. Equation (2.28) may be viewed as an equation to be solved for μ^*. It is not expected to have a solution unless the y in $\alpha_s=s^y$ is properly chosen. This seems reasonable if we consider the case $s\to\infty$. We expect that all factors of s (and hence y) must delicately balance to achieve (2.28). In some sense (2.28) is an "eigenvalue equation" for the eigenvalue y and eigenvector μ^*. Of course, (2.28) is not a linear equation. We have no theorem so far to tell us whether (2.28) has a discrete, or a continuous set of solutions, or even any solution at all. For the moment, we simply assume that there is at least one solution. We shall concentrate on a particular one with a definite y. We define the quantity η for this y:

$$y=1-\tfrac{1}{2}\eta, \tag{2.29}$$

then

$$\alpha_s=s^{1-\frac{1}{2}\eta}. \tag{2.30}$$

We shall identify η as a critical exponent later. Equation (2.23) now takes the form

$$G(k,\mu)=s^{2-\eta}G(sk,R_s\mu),\quad k<\Lambda/s. \tag{2.31}$$

This formula will be used very often later.

More general correlation functions can be defined. For example, let

$$\begin{aligned}G_{i_1i_2\cdots i_m}(k_2,k_3\cdots k_m,\mu)&\\ \equiv\int d^dx_2d^dx_3\cdots d^dx_m\exp(-ik_2\cdot x_2-\cdots-ik_m\cdot x_m)&\\ \times\langle\phi_{i_1}(0)\phi_{i_2}(x_2)\cdots\phi_{i_m}(x_m)\rangle_P,&\\ =L^{(d/2)m-d}\langle\phi_{i_1k_1}\phi_{i_2k_2}\cdots\phi_{i_mk_m}\rangle_P,&\end{aligned} \tag{2.32}$$

where $k_1=-k_2-k_3-\cdots-k_m$ and none of the subsums of the k's is zero. It is easy to generalize (2.31) to

$$G_{i_1\cdots i_m}(k_2\cdots k_m,\mu)=s^{(m/2)(d+2-\eta)-d}G_{i_1\cdots i_m}(sk_2\cdots sk_m,R_s\mu), \tag{2.33}$$

provided that $k_1, k_2, \cdots, k_m<\Lambda/s$.

F. R_s as a Refined Scale Transformation

The transformation R_s can be viewed as a scale transformation. It tells how coupling parameters change when the system is shrunk by a factor s. However, the multiple integral and the determination of α_s by a fixed point equation make R_s very different from a naive change of scale. The multiple integral in (2.21) is necessary to keep the cutoff Λ fixed under R_s, i.e., it changes Λ to Λ/s and then lets the scale change bring Λ/s back to Λ. This is an extremely important point. The coupling parameters are defined with respect to a definite Λ. To compare two sets of coupling parameters, we must make sure that they are defined

with respect to the same cutoff. Therefore, to define a sensible scale transformation, it is necessary to keep Λ fixed. The multiple integral is an unambiguous way. Thus, R_s can be viewed as a refined scale transformation keeping the cutoff fixed.

As was mentioned below (2.27), the quantity y can be interpreted as the dimension of ϕ_k in units of length. In (2.29) we have chosen $y=1-\frac{1}{2}\eta$ to be an interaction-dependent quantity based on the fixed-point equation (2.28). Thus, the concept of dimension of a random variable under our refined scale transformation becomes an interaction-dependent concept.

G. Wilson's Recursion Formula

The first explicit calculation with the renormalization group as defined by (2.21) was carried out approximately by numerical means by Wilson (1971). We shall outline some basic features of his method of calculation.

It will be evident later that the transformation of interest is R_s with large s. The usefulness of the renormalization group is not affected if we restrict s to

$$s=2^l, \qquad l=0, 1, 2, 3, \cdots \tag{2.34}$$

so that R_s is just applying R_2 l times:

$$R_s=(R_2)^l. \tag{2.35}$$

One then works out $R_2\mu$ for a general μ. The result is the recursion formula of Wilson. The renormalization group is then obtained by repeated applications of the recursion formula.

Note that regarding R_s as R_2 repeated l times is not just a change of terminology. It exhibits the two distinctive features of R_s of large s, i.e., first the transformation R_2, and second, the *repetitions*. It is the large number of repetitions that will be directly related to the singularities in critical behavior. R_2 is the "generator" of the renormalization group.

Separating the tasks of obtaining and of repeating R_2 also allows some flexibility in computing and making approximations. For example, Wilson's approximate recursion formula for R_2 was obtained by using "wave packet variables" as integration variables in the multiple integral of (2.21). We shall briefly sketch the basic idea which can be generalized for other applications.

The random variable ϕ_k denotes the fluctuating amplitude of a plane wave configuration $\exp(ik\cdot x)$, which is spread over the whole volume. We expect $\mathcal{H}$ to be simpler when it is written in terms of more "localized" fluctuations because the interactions are assumed to be short range. Thus, it should be useful to introduce the new variables (wave packet variables)

$$\bar{\phi}(x_m)\equiv L^{-d/2}\sum_{\frac{1}{2}\Lambda<k<\Lambda}\phi_k\exp(ik\cdot x_m), \tag{2.36}$$

where the points x_m form a lattice. The spacing between lattice points is such that the total number of variables $\bar{\phi}(x_m)$ is the same as the number of ϕ_k's with k in the shell $\frac{1}{2}\Lambda<k<\Lambda$. The new variable $\bar{\phi}(x_m)$ represents the fluctuating amplitude of the wavepacket configuration

$$L^{-d}\sum_{\frac{1}{2}\Lambda<k<\Lambda}\exp[ik\cdot(x-x_m)] \tag{2.37}$$

centered around x_m. This is the "most localized" configuration one can construct by superimposing plane waves of wave vectors in the shell $\frac{1}{2}\Lambda<k<\Lambda$. By smoothing the wavepacket and using $\bar{\phi}(x_m)$ as integration variables in (2.21), Wilson worked out an approximate formula for $R_2\mu$, which is suitable for numerical work and also as a basis for further approximations.

H. Technical Remarks

1. *Smoothed Cutoff*

We add a technical note, which will be important in practical calculations, although it happens to be unimportant for our discussion of idealized cases.

The multiple integral in (2.21) implies a sharp cutoff in k space. That is to say that for k immediately below Λ/s, ϕ_k is not integrated but that it would be integrated if k is immediately above Λ/s. This sharp cutoff leads to undesirable mathematical features such as nonanalytic behavior of the u_{2m}'s as functions of the wave vectors. One also expects oscillating tails in the new coupling parameters of $\mathcal{H}'$ in the coordinate representation. This is analogous to the Friedel oscillation, which comes from the sharp Fermi surface, in the theory of Fermi gases. However, unlike the Friedel oscillation, these undesirable features here are of purely mathematical origin and will lead to no important consequences. They simply introduce complications in intermediate steps of the calculation. It is desirable to remove the sharp cutoff by making the transition from "integrated" to "unintegrated" smooth. This can be done [Wilson and Kogut (1972)], but it is too complicated to explain here. In the graphical representation to be introduced later this can be done easily. What we want to point out here is that the fixed point μ^* will depend on how the cutoff is effected.

2. *Another Technical Remark*

Because the spin $\phi(x)$ is a real quantity, its Fourier transform ϕ_k satisfies the relation

$$\phi_{ik}=\phi_{i-k}{}^*. \tag{2.38}$$

Here the * means complex conjugate, of course. Therefore ϕ_{i-k} cannot be regarded as a different random variable from ϕ_{ik}. What we should do is to use $\mathrm{Re}\,\phi_{ik}$ and $\mathrm{Im}\,\phi_{ik}$ as two real random variables and realize that

$$\int d\phi_{ik}\int d\phi_{i-k}=\int_{-\infty}^{\infty}d(\mathrm{Re}\,\phi_{ik})\int_{-\infty}^{\infty}d(\mathrm{Im}\,\phi_{ik}). \tag{2.39}$$

Thus the integrals in (2.21) must be taken a pair at a time.

I. A Reminder

We want to emphasize that so far we have only defined transformations R_s in the parameter space. We have not solved any problems. It is not even clear that such transformations are relevent in any way to physics, not to mention critical phenomena. The definition of R_s is separated from the concept of averages, above or below critical point, energy, temperature, etc. R_s simply takes one point in the parameter space to another.

III. BEHAVIOR AT LARGE s AND CRITICAL EXPONENTS

Let $\mu(T)$ be the point in the parameter space representing the canonical ensemble at temperature T. Namely, $\mu(T)$ represents the physical probability distribution. We shall argue that, if T is very close to T_c, $R_s\mu(T)$ will become close to the fixed point μ^* for very large s. Critical behaviors, in particular the critical exponents, will then be related to the properties of R_s operating near the fixed point. Near the fixed point, we can imagine a "linearized" R_s. Since R_s, as we have defined it, is a very complicated nonlinear transformation, its qualitative features are not at all easy to see. However, once it is linearized, our experience with linear operators and vector spaces will help us to make a lot of guesses. In this section our discussion will still be formal. Some of the statements will be unsubstantiated. Illustration by explicit calculation will come in Secs. IV and VI.

A. The Linearized Equation

If μ is near μ^*, we write formally

$$\mu=\mu^*+\delta\mu, \tag{3.1}$$

where $\delta\mu$ is small in some sense. The equation $\mu'=R_s\mu$ can be written as

$$\delta\mu'=R_s^L\delta\mu \tag{3.2}$$

since $R_s\mu^*=\mu^*$, $\mu'\equiv\mu^*+\delta\mu'$. R_s^L becomes a linear operator when $O((\delta\mu)^2)$ terms are dropped in calculating $\delta\mu'$ from (3.2). In principle, at least, we can construct a matrix to represent R_s^L in (3.2); and can determine the eigenvalues and eigenvectors of this matrix. Suppose that the eigenvalues are found to be $\lambda_j(s)$ and the corresponding eigenvectors to be e_j, $j=1, 2, 3, \cdots, \infty$. We label the eigenvalues in the order $\lambda_1\geq\lambda_2\geq\lambda_3\cdots$. Note that since $R_sR_{s'}e_j=R_{ss'}e_j$, we have

$$\lambda_j(s)\lambda_j(s')=\lambda_j(ss')$$

$$\therefore\lambda_j(s)=s^{y_j}, \tag{3.3}$$

where y_j are constants and $y_1\geq y_2\geq y_3\cdots$, since $s\geq 1$. We write $\delta\mu$ as a linear combination of the eigenvectors e_j:

$$\delta\mu=\sum_j t_je_j; \tag{3.4}$$

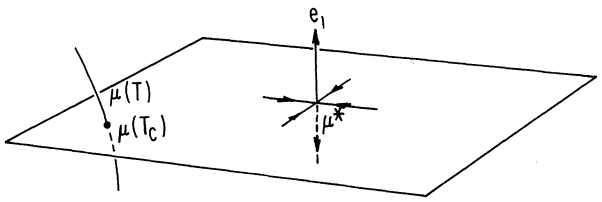

FIG. 1. Qualitative picture of a critical surface and a fixed point μ^* in the parameter space. The arrows point in directions of motion of $R_s\mu$ as s increases. The trajectory on the left is $\mu(T)$ for a continuous range of T, and $\mu(T_c)$ is the intersection of the trajectory and the critical surface.

then from (3.2) and (3.3), we have

$$\delta\mu'=\sum_j t_js^{y_j}e_j. \tag{3.5}$$

Apparently, we have made no progress since we do not know y_j nor e_j. But simplicity appears if it turns out that only $y_1>0$, all other y_j's are negative. In this case we have

$$\delta\mu'=R_s^L\delta\mu=t_1s^{y_1}e_1+O(s^{y_2}) \tag{3.6}$$

if s is so large that the first term dominates but $t_1s^{y_1}$ is still small enough so that the linear approximation for R_s is valid. If $t_1=0$ to start with, then $R_s^L\delta\mu\to 0$ as s increases; i.e., μ will be "pushed" toward the fixed point by R_s. Wilson calls t_1 a "relevant" variable and the other t_j's "irrelevant."

We can imagine that the eigenvectors e_j span the linear vector space which is the neighborhood of μ^*. The subspace defined by $t_1=0$ will be called the "critical surface." Points on the critical surface will be pushed to the fixed point by R_s, and points not on the critical surface will be pushed toward e_1 but away from the fixed point as (3.6) indicates. [See Fig. 1.]

The linear approximation for R_s is expected to break down when μ, μ' are not very close to μ^*. But we expect the general picture of a critical surface and the approach to the e_1 axis of $R_s\mu$ for large s to remain valid.

B. Critical Exponents and the Correlation Length

So far no physical concept has appeared in our discussion of the renormalization group. R_s simply transforms one probability distribution to another in a peculiar way. Now we shall examine the effect of R_s on the probability distribution (2.11), which describes fluctuations in a physical system at a definite temperature. This particular probability distribution is represented by a certain point $\mu(T)$ in the parameter space. This point corresponds to a set of coupling parameters which depend on the temperature T. They must be a smooth function of T. Because we have integrated out $\phi_{k'}$ with $k'>\Lambda$ in the microscopic Hamiltonian [see (2.11)], $H(\Lambda)$ would depend on T also. It is important to note that the integrations are over $\phi_{k'}$ with large k' and that we would not expect any singular temperature dependence of $H(\Lambda)$ due to such integrals. If we vary T continuously, we would trace out a trajectory in the

parameter space. This trajectory should be very smooth, and hits the critical surface at a special temperature T_c as shown in Fig. 1. At a temperature T which is very close to T_c and assuming $\mu(T)$ to be close to μ^*, the distance from $\mu(T)$ to the critical surface, which is t_1, is then proportional to $T-T_c$; i.e., t_1 is a function of T and can be expanded as

$$t_1(T)=A(T-T_c)+B(T-T_c)^2+\cdots.$$

We assume that $A\neq 0$. Let us assume that $\mu(T)$ is close to μ^*. If we write $\mu(T)=\mu^*+\delta\mu(T)$, then (3.6) reads

$$R_s^L\delta\mu(T)=A(T-T_c)s^{1/\nu}e_1+O(s^{y_2}), \tag{3.7}$$

where we have defined ν by

$$1/\nu=y_1. \tag{3.8}$$

Applying (3.7) to (2.31), we obtain, for large s,

$$G(k,\mu(T))=s^{2-\eta}[G(sk,\mu^*+A(T-T_c)s^{1/\nu}e_1+O(s^{y_2}))]. \tag{3.9}$$

Consider first the case $T=T_c$. Since s is arbitrary, we choose it to be proportional to $1/k$, say $s=\Lambda/2k$. We then get from (3.9)

$$G(k,\mu(T_c))=k^{-2+\eta}(\Lambda/2)^{2-\eta}[G(\Lambda/2,\mu^*)+O((\Lambda/2k)^{y_2})]. \tag{3.10}$$

In the limit of small k, this means

$$G(k,\mu(T_c))\propto k^{-2+\eta}, \tag{3.11}$$

which is the equation defining the critical exponent η. Thus, the critical exponent η is related to the fixed point equation (2.28). The power law (3.11) for $G(k)$ at T_c is seen as a consequence of the fact that $R_s\mu(T_c)$ approaches μ^* for large s. How small must k be in order for (3.11) to be a good approximation? Equation (3.10) says that $(2k/\Lambda)^{-y_2}$ must be small, much smaller than $1/2$, say; i.e.,

$$2k/\Lambda\ll 2^{1/y_2}. \tag{3.12}$$

Equation (3.12) is an estimate of the size of the critical region *in k space*, namely the region in which (3.11) holds. This size therefore strongly depends on y_2. Recall that s^{y_2} is the second largest eigenvalue of R_s in the linear approximation, and y_2 is assumed to be negative.

Now we consider the case $T-T_c>0$, $k=0$. We choose $s=t_1^{-\nu}$. Here we write t_1 for $A(T-T_c)$. Equation (3.9) gives

$$G(0,\mu(T))=t_1^{-(2-\eta)\nu}[G(0,\mu^*+e_1)+O(t_1^{-\nu y_2})]. \tag{3.13}$$

In the limit of small t_1; i.e., small $T-T_c$, we have

$$G(0,\mu(T))\propto(T-T_c)^{-\gamma}, \tag{3.14}$$

$$\gamma=\nu(2-\eta). \tag{3.15}$$

Equation (3.14) is the definition of the critical exponent γ and Eq. (3.15) is a "scaling law" relating the exponents γ, η, and ν. Equation (3.14) holds when $t_1^{-\nu y_2}$ is much smaller than order unity, say $1/2$, as (3.13) indicates. This means that

$$t_1\ll 2^{1/\nu y_2}. \tag{3.16}$$

In a way similar to (3.12), (3.16) estimates the size of the critical region in $T-T_c$. Equations (3.12) and (3.16) are oversimplified to exhibit the role of y_2. Many other parameters will generally enter in determining the size of the critical region. In other words, instead of $2^{1/y_2}$, we should have a complicated model dependent constant raised to the power $1/y_2$. The relevant question to answer for determining the size of the critical region is how large s must be so that $R_s\mu(T)$ is well approximated by $\mu^*+t_1s^{1/\nu}e_1$. Intuitively, we expect that the farther away $\mu(T)$ is from μ^*, the larger an s is required, and hence the smaller the critical region becomes. We shall have an opportunity to examine this point more explicitly later.

We now define the quantity ξ as

$$\xi=|t_1|^{-\nu}. \tag{3.17}$$

which we shall call the "correlation length." Then (3.7) reads

$$R_s^L\delta\mu=(s/\xi)^{1/\nu}e_1+O(s^{y_2}). \tag{3.18}$$

The effect of R_s is thus to decrease the correlation length by a factor s. If we ignore the $O(s^{y_2})$ term, we would then arrive at the scaling hypothesis discussed in the Introduction. Thus the scaling hypothesis is valid if R_s, in its linear approximation near μ^*, is dominated by one eigenvalue for large s.

What about the case where $T-T_c<0$? In this case, $t_1<0$, we can simply set $s=(-t_1)^{-\nu}$ and replace (3.13) by

$$G(0,\mu(T))=(-t_1)^{-\gamma}[G(0,\mu^*-e_1)+O((-t_1)^{-\nu y_2})]. \tag{3.19}$$

This is a correct statement but, in this case, it contains no information because $G(0,\mu)=\infty$ for $t_1<0$ [see Brezin, Wallace, and Wilson (1973)]. We shall not discuss G for the case $t_1<0$ in this article.

The assumption that μ must be near μ^* can in fact be relaxed. The critical surface can be taken as a surface extending away from μ^*. Any μ on this surface has the property that

$$\lim_{s\to\infty}R_s\mu=\mu^*. \tag{3.20}$$

For s large enough, $R_s\mu$ will be in the neighborhood of μ^*, and the linear approximation will then apply. It is clear that if μ is not close to μ^* but is very close to the critical surface, then there is some range of s for which $R_s\mu$ is not far away from μ^*. There is no need to find all the eigenvectors and eigenvalues of R_s^L. All we need to know is $1/\nu$ and y_2, which should be regarded as specifying the leading s dependence of $R_s\mu$ for large s.

Therefore, the qualitative conclusion above should hold for $\mu(T)$ close to the critical surface; i.e. for $T-T_c$ very small, but not necessarily close to the fixed point.

C. The Free Energy

So far we have focused our attention on the correlation function of a small wave number. Our conclusions rely heavily on formula (2.31). Similarly, using (2.33), we can obtain critical behaviors for more complicated correlation functions of type (2.32). Note that these correlation functions involve only averages of ϕ_{ik}'s of small k. There are quantities of interest which involve ϕ_{ik}'s of large k. For example, the free energy cannot be expressed as an average of the type (2.32). Variables ϕ_{ik} of all k are involved. What can we say about such quantities from what we know about R_s? The following study of the free energy will give a qualitative answer.

First, we want to derive a formula similar to (2.31) for the free energy. Let us start from the beginning.

The free energy per unit volume $F(T)$ is defined by

$$\exp\left[-L^d F(T)/T\right]=\prod_{0<k'<\Lambda}\int d\phi_{k'}\exp\left[-H(\Lambda)/T\right]. \tag{3.21}$$

For simplicity of notation, we shall not write out the component indices i explicitly. Clearly an additive constant in $H(\Lambda)$ will make a difference in F. To apply the renormalization group to the study of the free energy, one must specify the additive constant in $\mathcal{H}$ so far ignored. We shall adopt the rule that the additive constant is *always written out explicitly* and that symbols such as $\mathcal{H}$, $\mathcal{H}'$ *will contain no additive constant*; i.e., $\mathcal{H}$, $\mathcal{H}'$ are zero if $\phi=0$. We now define $\mathfrak{F}=\mathfrak{F}(\mu)$ by

$$\exp\left(-L^d\mathfrak{F}\right)=\prod_{0<k<\Lambda}\int d\phi_k\exp\left(-\mathcal{H}\right). \tag{3.22}$$

Similarly we define $\mathfrak{F}'=\mathfrak{F}(\mu')$ by replacing in (3.22) $\mathcal{H}$ by $\mathcal{H}'$, $L^d\mathfrak{F}$ by $L'^d\mathfrak{F}'$ and keeping in mind that the density of points in k space over which the product $\prod$ runs is changed to $L'^d(2\pi)^{-d}$. To relate $\mathfrak{F}'$ to $\mathfrak{F}$, we separate the multiple integral in (3.22) into two and write

$$\exp\left(-\mathfrak{F}L^d\right)=\prod_{0<k<\Lambda/s}\int d\phi_k\prod_{\Lambda/s<k'<\Lambda}\int d\phi_{k'}\exp\left(-\mathcal{H}\right). \tag{3.23}$$

We then make the substitution $\phi_k\rightarrow\alpha_s\phi_{sk}$ for the second (the left) set of variables and obtain

$$\begin{aligned}\exp\left(-\mathfrak{F}L^d\right)&=\prod_{0<sk<\Lambda}\int d\phi_{sk}\alpha_s\\&\quad\times\Big[\prod_{\Lambda/s<k'<\Lambda}\int d\phi_{k'}\exp\left(-\mathcal{H}\right)\Big]_{\phi_k\rightarrow\alpha_s\phi_{sk}},\\&=\prod_{0<k<\Lambda}\int d\phi_k\exp\left(-\mathcal{H}'\right)\\&\qquad\times\exp\left[-L^d(A+A_0)\right],\\&=\exp\left[-L'^d\mathfrak{F}'-(A+A_0)L^d\right].\end{aligned} \tag{3.24}$$

We have applied the definition (2.21) for $\exp(-\mathcal{H}')$. The constants A and A_0 are defined by

$$\exp\left(-L^dA\right)=\Big[\prod_{\Lambda/s<k'<\Lambda}\int d\phi_{k'}\exp\left(-\mathcal{H}\right)\Big]_{\phi_k=0}, \tag{3.25}$$

where ϕ_k, $0<k<\Lambda/s$, the unintegrated variables, are set to zero, and

$$\exp\left(-L^dA_0\right)=\prod_{0<k<\Lambda/s}\alpha_s{}^n;$$

i.e.,

$$A_0=-L^{-d}\sum_{k<\Lambda/s}n(1-\tfrac{1}{2}\eta)\ln s \tag{3.26}$$

since $\alpha_s=s^{1-\eta/2}$. The additive constant A would be just $\mathfrak{F}$ if all ϕ_k with $0<k<\Lambda/s$ were set to zero. A_0 is to compensate for the change of the size of the phase space produced by the substitution $\phi\rightarrow\alpha_s\phi$. We have therefore

$$\mathfrak{F}(\mu)=s^{-d}\mathfrak{F}(\mu')+A+A_0 \tag{3.27}$$

from (3.24). Now if $H(\Lambda)/T$ is represented by $\mu(T)$, the free energy is

$$F(T)=\mathfrak{F}(\mu(T))T. \tag{3.28}$$

We obtain from (3.27)

$$\begin{aligned}\mathfrak{F}(\mu(T))-\mathfrak{F}(\mu(T_c))&=s^{-d}\left[\mathfrak{F}(\mu'(T))-\mathfrak{F}(\mu'(T_c))\right]\\&\quad+A(T)-A(T_c).\end{aligned} \tag{3.29}$$

For large s, $\mu'(T_c)$ approaches the fixed point μ^*. If $T-T_c$ is very small, we choose

$$s=|t_1|^{-\nu}\propto|T-T_c|^{-\nu}, \tag{3.30}$$

as was done in (3.13) and obtain from (3.29)

$$\begin{aligned}&\mathfrak{F}(\mu(T))-\mathfrak{F}(\mu(T_c))\\&\quad=|t_1|^{\nu d}\left[\mathfrak{F}(\mu^*\pm e_1)-\mathfrak{F}(\mu^*)+O(|t_1|^{-\nu y_2})\right]\\&\qquad+(A(T)-A(T_c))_{s=|t_1|^{-\nu}},\end{aligned} \tag{3.31}$$

where, in the argument of $\mathfrak{F}(\mu^*\pm e_1)$, the $+$ and $-$ signs correspond to the cases $t_1>0$ and $t_1<0$, respectively. In the small $T-T_c$ limit, we have

$$F(T)-F(T_c)\propto|T-T_c|^{\nu d}+\text{"less singular terms"} \tag{3.32}$$

provided that the last two terms of (3.31) are truly less singular. Note that $\mathfrak{F}(\mu^*+e_1)$ is expected to be different from $\mathfrak{F}(\mu^*-e_1)$. Therefore the proportionality constant in front of $|T-T_c|^{\nu d}$ in (3.32) for $T>T_c$ is different from that for $T<T_c$.

Equations (3.31) and (3.32) are similar to (3.13) and (3.14). Here we cannot say very much about the less singular terms. We shall assume that they are indeed less singular. This assumption is correct in many cases but incorrect in some. Qualitatively speaking, the term $|T-T_c|^{\nu d}$ in (3.32) comes directly from the spin fluctuations with small k's (smaller than ξ^{-1}) and the other term comes from ϕ_k with large k's (larger than ξ^{-1}).

The specific heat exponent α is defined by

$$-\partial^2 F/\partial T^2 \propto |\, T-T_c\,|^{-\alpha}+\text{less singular terms.} \quad (3.33)$$

From (3.32), we then obtain the scaling law

$$\nu d=2-\alpha. \quad (3.34)$$

Besides the free energy, there are many other quantities of interest which involve ϕ_k with large k's in an essential way as well as small k's. One must carefully derive relations like (3.29) from the beginning before applying scaling or renormalization group arguments.

D. Nonzero Magnetization

The average value of $\phi_i(x)$ is zero as a result of the assumed rotation invariance in the n-dimensional spin vector space of the probability distribution and the assumption that μ is above the critical surface. This average value becomes nonzero when an external field H [see (2.9)] is turned on. It remains nonzero also when μ lies below the critical surface even when H is turned off. In the latter case, we have the rather striking phenomenon that a rotationally symmetric probability distribution produces apparently nonsymmetric average values. This is, of course, the most conspicuous feature of a phase transition. In a manner similar to the above conclusions concerning the correlation function and the free energy, we can also say something about the average of ϕ.

In defining the parameter space [see (2.12) and (2.16)], odd powers of ϕ were excluded. Now we introduce one more parameter H by adding to $\mathcal{H}$ a term

$$H\int d^dx\phi_1(x)=HL^{d/2}\phi_{10}, \quad (3.35)$$

where ϕ_{10} means $(\phi_{1k})_{k=0}$. The parameter H can be identified as proportional to a uniform external field in the 1 direction. (It should not be confused with a Hamiltonian.) It is easy to find out how H changes under R_s through (2.21). Since ϕ_{10} is never involved in the multiple integral, the only thing happening is the replacement $\phi_{10}\to\alpha_s\phi_{10}$. Thus, in $\mathcal{H}'$, there appears a term

$$HL^{d/2}s^{1-\eta/2}\phi_{10}=H'L'^{(d/2)}\phi_{10}, \quad (3.36)$$

$$H'=s^{\frac{1}{2}(d-\eta)+1}H. \quad (3.37)$$

Recall that $\alpha_s=s^{1-\eta/2}$, $sL'=L$. We can therefore write

$$(H',\mu')=(H',R_s\mu), \quad (3.38)$$

with H' given by (3.37) and $R_s\mu$ defined as previously, as the renormalization group transformation in the extended parameter space. The average "magnetization" M is given by

$$M(H,\mu)=\langle\phi_1(x)\rangle_P=L^{-d/2}\langle\phi_{10}\rangle_P, \quad (3.39)$$

where P of course denotes the probability distribution represented by (H,μ). We know that

$$\langle\phi_{10}\rangle_P=s^{1-\eta/2}\langle\phi_{10}\rangle_{P'}, \quad (3.40)$$

where P' stands for the probability distribution represented by (H',μ'). [If (3.40) is not obvious, please go back to the three trivial facts discussed at the beginning of Sec. II. See (2.5) in particular.] Substituting (3.40) in (3.39), we obtain an equation analogous to (2.31):

$$\begin{aligned} M(H,\mu)&=L'^{-d/2}s^{-d/2}s^{1-\eta/2}\langle\phi_{10}\rangle_{P'}\\ &=M(H',\mu')s^{-\frac{1}{2}(d+\eta)+1}\\ &=M(Hs^{\frac{1}{2}(d-\eta)+1},\mu')s^{-\frac{1}{2}(d+\eta)+1}. \end{aligned} \quad (3.41)$$

Before we proceed further, let us emphasize that as long as M and H are uniform, the renormalization group transformation $\mu'=R_s\mu$ discussed previously is not affected, regardless whether μ is above, on, or below the critical surface.

If $\mu=\mu(T_c)$ is a point on the critical surface, μ' will approach μ^* for large s. If H is small enough (i.e., weak external field), we can choose

$$s=H^{-[\frac{1}{2}(d-\eta)+1]^{-1}} \quad (3.42)$$

so that (3.41) becomes

$$M(H,\mu(T_c))=H^{1/\delta}M\{1,\mu^*+O(H^{-y_2[\frac{1}{2}(d-\eta)+1]^{-1}})\}, \quad (3.43)$$

where

$$\delta=(d+2-\eta)/(d-2+\eta). \quad (3.44)$$

In the limit of small H, we have

$$M\propto H^{1/\delta} \quad (3.45)$$

which is the equation defining the exponent δ.

If $H=0$, and μ is below the critical surface, we choose $s=|\,t_1\,|^{-\nu}$ and obtain from (3.41)

$$M(\mu)=|\,t_1\,|^{\beta}M(\mu^*-e_1+O(|\,t_1\,|^{-\nu y_2})), \quad (3.46)$$

$$\beta=\tfrac{1}{2}\nu(d-2+\eta). \quad (3.47)$$

The exponent β is defined by $M\propto|\,T-T_c\,|^{\beta}$ in the limit of small $|\,T-T_c\,|$ below T_c.

At this point we would like to remind the reader that R_s only transforms the parameters, and that all the above conclusions are consequences of the assumed properties of R_s for large s near the fixed point. The over-all pattern is governed by three exponents in the above discussion, namely, η, $1/\nu\equiv y_1$, and y_2. If $O(s^{y_2})$ is neglected, the results we have are just those obtained from the scaling hypothesis. Details of $\mu(T)$ do not appear in the exponents, and this fact is called 'universality of exponents.'

Just by looking at the renormalization group, we of course cannot tell what average values the probability distribution will produce, since R_s involves no attempt to calculate any average value. In particular, we cannot tell if M is zero or not. In (3.46), we simply *assume* that M is not zero if $t_1<0$.

E. Remarks on the Status of Qualitative Conclusions

Let us review the basis of all the qualitative conclusions about critical behavior obtained so far.

We have the picture of a parameter space in which there is a fixed point μ^* of R_s. The fixed point sits on

a critical surface. R_s drives any point on this surface toward μ^*. For a point not on the critical surface but sufficiently close to it, R_s will first drive it toward μ^* and then drive it away in a direction which we call e_1. This picture is purely mathematical; i.e., no reference to physics is made. If the probability distribution describing a real physical system happens to be represented by a point very close to or on the critical surface, then this purely mathematical picture becomes useful in explaining critical phenomena in this system, as we have shown with the aid of identities such as (2.31) or (3.41). Therefore, the important question is whether this mathematical picture actually exists or is just a fiction. We have given no proof that it exists. In fact, there is no rigorous theorem telling us under what condition such a picture can emerge. There are cases, as we shall illustrate in the following sections, where it turns out to be real. In fact, such a picture is believed to be only one of many possible pictures. It may happen that, for example, y_2, as well as y_1, is positive. Such cases are useful in explaining tri-critical phenomena [Riedel and Wegner (1972)]. It might happen that there is more than one fixed point, or that there are important complex eigenvalues for R_s in the linear approximation. Different possibilities in the behavior of R_s are expected to be consequences of different symmetry restrictions and other features of the parameter space. It is desirable to have more rigorous work done so as to classify various possibilities. The difficulty is in the mathematical complication, not in the principle. In principle, the renormalization group is well defined and can always be carried out approximately by numerical means.

IV. THE RENORMALIZATION GROUP IN THE LARGE n CASE (SPHERICAL MODEL)

Having gone through the formal definitions and qualitative conclusions, it is desirable to see explicitly in an example what R_s really looks like and to verify that the fixed point and all that actually emerge from calculations. In this section, we study the case of large n and work out everything. The large n limit is directly related to the spherical model studied by Stanley (1968), and by Berlin and Kac (1952). Discussions and references on the spherical model can be found in Stanley's book (1971). Here we are interested in how the renormalization group looks explicitly. No previous knowledge concerning the spherical model is assumed here.

The major mathematical task is to evaluate the multiple integral in (2.21), which defines R_s. For large n, the integrand turns out to have a sharp maximum. The integral can be approximately evaluated by locating the maximum. Let us digress briefly on this point. Consider an integral

$$I=\int_{-\infty}^{\infty}\exp\left[-\lambda f(x)\right]dx,$$

where $f(x)$ is assumed to be real and bounded below. We also assume that $f(x)\to\infty$ as $|x|\to\infty$. Then the contribution to I comes dominately from near the maximum of the integrand which is very sharp for large λ. The location $\bar{x}$ of the maximum is obtained by setting the derivative of f to zero:

$$f'(\bar{x})=0.$$

Then, near the maximum of the integrand we have

$$f(x)\approx f(\bar{x})+\tfrac{1}{2}f''(\bar{x})\,(x-\bar{x})^2$$

and

$$\begin{aligned}I&\approx\exp\left[-\lambda f(\bar{x})\right]\int_{-\infty}^{\infty}\exp\left[-\tfrac{1}{2}\lambda(x-\bar{x})^2f''(\bar{x})\right]dx,\\&=\exp\left[-\lambda f(\bar{x})\right](2\pi/\lambda f''(\bar{x}))^{1/2},\\&=\exp\left[-(\lambda f(\bar{x})+O(\ln\lambda))\right].\end{aligned}$$

Thus, to find the logarithm of I to the leading order in λ, all we have to do is to find $\bar{x}$ and the approximate answer for $\ln I$ is simply $-\lambda f(\bar{x})$. The error is of $O(\ln\lambda)$. This procedure is the same as those in the methods of stationary phase and steepest descent. In our case, the large parameter corresponding to λ will be n. Instead of a single integral, we shall have a multiple integral. The determination of the location of the sharp maximum will lead to a nonlinear equation familiar in Hartree approximations in many-body theory. R_s can then be obtained by solving such an equation.

The fixed point and related objects discussed in Sec. III will then follow. We shall find that the mathematical picture assumed in Sec. III does emerge. We shall see that $\eta=0$, $y_1=d-2$, $y_2=d-4$ for $2<d<4$; and $\eta=0$, $y_1=2$, $y_2=4-d$ for $d>4$. The fixed point for $d\geq 4$ will be trivial but for $2<d<4$ it will be more complicated. Our procedure will not be applicable if $d\leq 2$. The results in this section have been reported by Ma (1973b, c).

What follows is some "straightforward but tedious algebra," which requires some patience to go through. For a reader who wants to see the answers first, we suggest that he go through the definition of the parameter space [(4.1)–(4.4)] first, then jump to (4.34) and (4.35), take them for granted, and read on. He should later return to what he skipped to find out what approximations have been made.

A. The Parameter Space

We shall not need the general parameter space defined by (2.16)–(2.19). We need a parameter space just a bit more complicated than that defined by the Landau–Ginzburg form (2.12) and (2.15). We write

$$P\propto\exp\left(-\mathcal{H}\right)$$

$$\mathcal{H}=\int d^dx\left(a(\nabla\phi)^2+U(\phi^2)\right),\qquad(4.1)$$

where $(\nabla\phi)^2$ and ϕ^2 are defined in (2.13) and $U(\phi^2)$ is a power series in ϕ^2. The first two terms are included in

(2.12). The parameter space is thus the space of

$$\mu=(a, t_0, u_4, u_6, \cdots). \tag{4.2}$$

The function U specifies all the parameters t_0, u_4, u_6, $\cdots$ which are coefficients of the powers of ϕ^2 in $U(\phi^2)$. The abbreviations (2.13) apply here. Since any additive constant in U would not be counted as a parameter, the derivative

$$dU(\phi^2)/d\phi^2 \equiv t(\phi^2) \tag{4.3}$$

specifies all the parameters t_0, u_4, u_6, $\cdots$. We can therefore simply write

$$\mu=(a, t). \tag{4.4}$$

The parameter space is the product space of the real numbers and the space of all power series. We require that P be integrable. This means that t must be such that $U(\phi^2)\to\infty$ as $\phi^2\to\infty$.

B. The Multiple Integral

The most important ingredient of the renormalization group is the multiple integral as given in (2.21). Unfortunately, to perform the integral is a very difficult task. We shall make approximations to simplify it.

We shall introduce the following notation to simplify writing

$$\Pi' \equiv \prod_{\Lambda/s<k'<\Lambda},$$
$$\sum' \equiv \sum_{\Lambda/s<k'<\Lambda},$$
$$\sum \equiv \sum_{k<\Lambda/s}. \tag{4.5}$$

In terms of the Fourier components ϕ_{ik}, we can separate $\phi_i(x)$ into two pieces; one piece is the sum over $\Lambda/s<k<\Lambda$, which we shall call $\phi_i'(x)$:

$$\phi_i'(x)=L^{-d/2}\sum' \phi_{ik'} \exp(ik'\cdot x), \tag{4.6}$$

and the rest, which contains ϕ_{ik} which will not be integrated, will simply be called $\phi_i(x)$ again (to avoid introducing more notation); i.e.,

$$\phi_i(x)\to\phi_i(x)+\phi_i'(x). \tag{4.7}$$

The gradient term in (4.1) becomes simply

$$\int d^dx a(\nabla\phi)^2+a\sum' k'^2 N_{k'}, \tag{4.8}$$

where

$$N_{k'}\equiv\tfrac{1}{2}\sum_{i=1}^{n}|\phi_{ik'}|^2. \tag{4.9}$$

We have, squaring (4.7) and summing over i,

$$\phi^2(x)\to\tfrac{1}{2}\sum_{i=1}^{n}(\phi_i^2(x)+\phi_i'^2(x)+2\phi_i(x)\phi_i'(x)) \tag{4.10}$$

$$\approx\phi^2(x)+\tfrac{1}{2}\sum_{i=1}^{n}\phi_i'^2(x). \tag{4.11}$$

We have dropped the last term in (4.10). This is permissible for very large n. Intuitively, the reason is the following. Since n is large, the first two terms, being sums of n positive quantities, are large. The last term, sometimes positive and sometimes negative, tends to cancel as we sum over i. It will be more evident later that the probability distribution sharply peaks at the average values of random variables owing to the large n. Dropping the last term of (4.10) makes an error of $O(1/n)$. We shall make one more approximation on the last term of (4.11):

$$\tfrac{1}{2}\sum_{i=1}^{n}\phi_i'^2(x)=\tfrac{1}{2}L^{-d}\sum_{k'}'\sum_{k''}'\sum_{i=1}^{n}\phi_{ik'}\phi_{ik''}^*\times\exp[i(k'-k'')\cdot x]$$
$$\approx L^{-d}\sum' N_{k'}, \tag{4.12}$$

where $N_{k'}$ is defined by (4.9). Namely, we keep only terms $k''=k'$. Again, the intuitive reason is that, when n is very large, the sum over absolute squares in $N_{k'}$ is expected to be much larger than the terms dropped. A more detailed examination would show that the approximations in (4.11) and (4.12) are equivalent. The qualitative picture is the following. We have a spin vector $\phi(x)$ of many components. While each component can fluctuate a great deal, the *length* of the spin vector ($2\phi^2$ is the square of this *length*) being a large quantity is expected to fluctuate around its mean value over a very small fraction of its mean value. The above approximations effectively regard ϕ^2 as a constant.

Now we can write the $\mathcal{H}$ in (4.1) as

$$\mathcal{H}=\int d^dx a(\nabla\phi)^2+\sum' ak'^2N_{k'}+\int d^dx U(L^{-d}\sum' N_{k'}+\phi^2). \tag{4.13}$$

All the variables $\phi_{ik'}$, $\Lambda/s<k'<\Lambda$, to be integrated over appear through $N_{k'}$. Therefore, we shall change the integration variables to $N_{k'}$. We write, following (2.39),

$$\prod_{i=1}^{n}\int d\phi_{ik'}\,d\phi_{i-k'}=\prod_{i=1}^{n}\int_{-\infty}^{\infty}d(\mathrm{Re}\,\phi_{ik'})\,d(\mathrm{Im}\,\phi_{ik'})$$
$$\propto\int_0^\infty d(2N_{k'})^{1/2}[(2N_{k'})^{1/2}]^{2n-1}$$
$$\propto\int_0^\infty dN_{k'}N_{k'}{}^{n-1}. \tag{4.14}$$

The proportionality constant in (4.14) can be ignored since it can only introduce an additive constant to $\mathcal{H}'$ in (2.21). Such additive constants play no role. Equation (4.14) follows from the fact that, by (4.9), $(2N_k)^{1/2}$ is the length of the $2n$ component vector ($\mathrm{Re}\,\phi_{1k}$, $\mathrm{Re}\,\phi_{2k}$, $\cdots$, $\mathrm{Re}\,\phi_{nk}$; $\mathrm{Im}\,\phi_{1k}$, $\cdots$, $\mathrm{Im}\,\phi_{nk}$). Then the multiple integral in (2.21) is

$$\Pi'\int_0^\infty dN_{k'}N_{k'}{}^{n-1}\exp(-\mathcal{H}). \tag{4.15}$$

In the product half of the points k' must be excluded

since each $dN_{k'}$ takes care of $d\phi_k$ and $d\phi_{-k}$, as (4.14) shows. This is just a technical point. All we need to fix up is a factor 1/2 somewhere later.

If n is very large, an approximate evaluation of the integral is not difficult. This is because, for large n, the factors $N_{k'}{}^{n-1}$ grow very rapidly for large $N_{k'}$, and meanwhile $\exp(-\mathcal{H})$ falls to zero very sharply. As a result, the integrand has a very sharp maximum and the contribution to the integral comes dominantly from the neighborhood of this sharp maximum. Let us write the integrand of (4.15) as

$$\prod' N_{k'}{}^{n-1} \exp(-\mathcal{H}) = \exp\{-\int d^dx[a(\nabla\phi)^2+W]\} \quad (4.16)$$

$$W \equiv L^{-d}\sum'[-\tfrac{1}{2}n \ln N_{k'}+ak'^2N_{k'}] + U(\rho'+\phi^2), \quad (4.17)$$

where

$$\rho' \equiv L^{-d}\sum' N_{k'}. \quad (4.18)$$

We have used (4.13) for $\mathcal{H}$. The factor $\frac{1}{2}$ in W follows from the remark below (4.15), and $(n-1)\ln N_{k'}$ is replaced by $n \ln N_{k'}$ since $n \gg 1$ is assumed. The maximum of the integrand can be located by setting $\partial W/\partial N_{k'}$ to zero for all k' in the shell $\Lambda/s<k'<\Lambda$. Let the solution of $\partial W/\partial N_{k'}=0$ be $\bar{N}_{k'}$. We obtain from (4.17)

$$(\partial W/\partial N_{k'})_{N=\bar{N}} = -(n/2\bar{N}_{k'})+ak'^2+t(\bar{\rho}'+\phi^2) = 0, \qquad \Lambda/s<k'<\Lambda, \quad (4.19)$$

where $\bar{\rho}'$ is obtained from (4.18) by setting $N_{k'}=\bar{N}_{k'}$ and t is the derivative of U by definition (4.3). Therefore, we have, from (4.19) and (4.18),

$$\bar{N}_{k'} = \tfrac{1}{2}n(ak'^2+t(\bar{\rho}'+\phi^2))^{-1}, \quad (4.20)$$

$$\bar{\rho}' = \tfrac{1}{2}nK_d\int_{\Lambda/s}^{\Lambda} dk'k'^{d-1}(ak'^2+t(\bar{\rho}'+\phi^2))^{-1}. \quad (4.21)$$

The sum over k' has been replaced by an integral

$$L^{-d}\sum' = \int[d^dk'/(2\pi)^d] = K_d\int_{\Lambda/s}^{\Lambda} dk'k'^{d-1}, \quad (4.22)$$

where K_d is $(2\pi)^{-d}$ times the area of a unit sphere in the d-dimensional Euclidean space:

$$K_d = 2^{-d+1}\pi^{-d/2}/\Gamma(\tfrac{1}{2}d). \quad (4.23)$$

This formula makes sense for d = positive integer. Here we simply use it as our definition of nonintegral dimensions.

Those who are familiar with many-body theory would recognize that (4.20) and (4.21) are the kind of equations which often occur in the Hartree approximation or the "self-consistent-field" approximation.

In view of our discussion at the beginning of this section, the integral (4.15) is just

$$\exp[-\int d^dx(a(\nabla\phi)^2+\bar{W})-C], \quad (4.24)$$

where C is of $O(\ln n)$ and is small compared to the other terms in the exponent which are of $O(n)$ and $\bar{W}$ is given by (4.17) with $N_{k'}$ replaced by $\bar{N}_{k'}$.

C. Formula for R_s

Now what is inside the bracket of (2.21), i.e., the multiple integral, is given by (4.24). We can read off $\mathcal{H}'$:

$$\mathcal{H}' = [\int d^dx(a(\nabla\phi)^2+\bar{W})]_{\phi_k\to\alpha_s\phi_{sk}}, \quad (4.25)$$

with $\alpha_s = s^{1-\eta/2}$. Then, we are able to write $\mathcal{H}'$ as

$$\mathcal{H}' = \int d^dx[a'(\nabla\phi)^2+U'(\phi^2)], \quad (4.26)$$

where the volume is now $L'^d \equiv s^{-d}L^d$. We are then able to differentiate U' to obtain

$$t'(\phi^2) = dU'/d\phi^2, \quad (4.27)$$

and get finally our $\mu' = (a', t') \equiv R_s\mu$. Let us go through these steps. The substitution $\phi_k \to \phi_{sk}s^{1-\eta/2}$ leads to

$$\phi_i(x) \to s^{-d/2+1-\eta/2}\phi_i(x/s), \qquad \phi^2 \to s^{2-d-\eta}\phi^2. \quad (4.28)$$

Since the new $\int d^dx$ means integrating over a smaller volume $L'^d = s^{-d}L^d$, we must make the replacement

$$\int d^dx \to s^d\int d^dx, \qquad x/s \to x. \quad (4.29)$$

It then follows that

$$\int d^dx a(\nabla\phi)^2 \to \int d^dx a s^{-\eta}(\nabla\phi)^2, \quad (4.30)$$

namely

$$a' = as^{-\eta}, \quad (4.31)$$

and that

$$U'(\phi^2) = s^d\bar{W}(\bar{N}_{k'}, s^{2-d-\eta}\phi^2), \quad (4.32)$$

Note that $\bar{N}_{k'}$ depends on ϕ^2 (now $s^{2-d-\eta}\phi^2$), via (4.20). In spite of this, the condition $\partial W/\partial N_{k'}=0$ at $N_{k'}=\bar{N}_{k'}$ allows us to obtain from (4.32)

$$t'(\phi^2) = dU'/d\phi^2, = (\partial U'/\partial\phi^2)_{\bar{N}_{k'}}, = s^{2-\eta}t(\bar{\rho}'+s^{2-d-\eta}\phi^2). \quad (4.33)$$

The last step comes from the fact that only the last term of (4.17) depends on ϕ^2 explicitly. Of course, t is the derivative of U. Therefore, given $\mu=(a, t)$, we can find $\mu'=(a', t')$ from (4.31) and (4.33) by solving (4.21). In order to have a fixed point with a finite and nonzero value of a, we must choose

$$\eta = 0 \quad (4.34)$$

as (4.31) implies. If η is chosen positive, and if there is a fixed point (a^*, t^*), then a^* must be zero. Some algebra will show that t^* would either be infinite or lead to a nonintegrable probability distribution. If η is chosen negative, a^* would be infinite. We shall keep to the interesting case of $\eta=0$. We now summarize our final

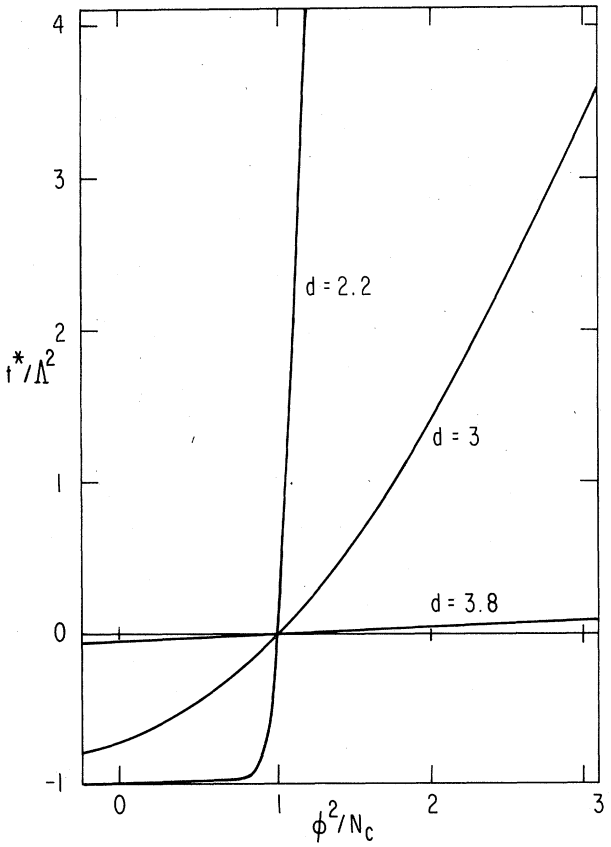

FIG. 2. Plot of $t^*(\phi^2)$ for $d=2.2$, 3, and 3.8. See (4.42). The unit N_c is given by (4.36) and depends on d.

formula for $R_s\mu=\mu'$:

$$t'(\phi^2)=s^2t(\bar{p}'+s^{2-d}\phi^2) \tag{4.35a}$$

$$\bar{p}'=\tfrac{1}{2}nK_d\int_{\Lambda/s}^{\Lambda}dkk^{d-1}/(k^2+t'/s^2). \tag{4.35b}$$

Since a remains always the same, we simply set it equal to 1. Generalization to $a\neq 1$ is trivial: just write ϕ for $(a)^{1/2}\phi$. Equation (4.35b) is just (4.21) with t written as t'/s^2 as permitted by (4.35a). These equations also show that R_s is a very complicated nonlinear transformation in this case.

D. The Critical Surface and the Fixed Point

We shall see how the critical surface, the fixed point, and other concepts discussed in Sec. III all come out of (4.35).

Since we have set the fixed parameter a to 1; i.e., now $\mu=(1, t)$, the parameter space is simply equivalent to the space of power series $t(\phi^2)$. We shall speak of t and μ as the same object.

If t is on the critical surface, t' will approach a fixed point t^* as $s\to\infty$. Let us determine the critical surface. First, we observe that, for $s\to\infty$, (4.35b) approaches the constant

$$N_c\equiv\tfrac{1}{2}nK_d\int_0^{\Lambda}dkk^{d-1}/k^2$$

$$=\tfrac{1}{2}nK_d\Lambda^{d-2}/(d-2), \tag{4.36}$$

and $s^{2-d}\phi^2\to 0$ as long as $d>2$. Thus, if t^* is finite, (4.35a) demands that

$$t_1\equiv t(N_c)=0. \tag{4.37}$$

This is therefore a necessary condition for t being on the critical surface. It also shows that if t is close to or on the critical surface, the argument of t in (4.35a) must be close to N_c. This suggests that we introduce the new notations N and ζ:

$$N\equiv\bar{p}'+s^{2-d}\phi^2\equiv(1+\zeta/s^2)N_c. \tag{4.38}$$

Writing $\bar{p}'$ in terms of ζ and ϕ^2 in (4.35b), and doing a little rearrangement, (4.35b) becomes

$$\phi^2/N_c=1+\zeta s^{d-4}-(d-2)\Lambda^{2-d}\int_{\Lambda}^{\Lambda s}dpp^{d-1}$$

$$\times[(t'+p^2)^{-1}-(p^2)^{-1}], \tag{4.39}$$

where $p\equiv k/s$, and we have divided the whole equation by N_cs^{2-d}. N_c is given by (4.36). Now, if $t_1=0$, then, for large but finite s, (4.35a) gives

$$t'=s^2t(N_c(1+\zeta/s^2))$$

$$=u_c\zeta+O(s^{-2}), \tag{4.40}$$

where

$$u_c\equiv N_c(dt(N)/dN)_{N=N_c} \tag{4.41}$$

is assumed to be positive.[1] Important results are evident now from (4.39) and (4.40).

If $2<d<4$, the term $\zeta s^{d-4}\to t^*/u_cs^{d-4}\to 0$ as $s\to\infty$ and (4.39) gives a nontrivial fixed point t^*:

$$\phi^2/N_c=1-(d-2)\Lambda^{2-d}\int_{\Lambda}^{\infty}dpp^{d-1}$$

$$\times[(t^*+p^2)^{-1}-(p^2)^{-1}]. \tag{4.42}$$

This equation defines t^* as a function of ϕ^2. Thus, we have shown that t will be driven to the fixed point t^* as long as $t_1=0$ and $u_c>0$. The condition $t_1=0$, given by (4.37), defines a hyperplane in the parameter space. The critical surface is just this hyperplane excluding the region in which $u_c\leq 0$. We have thus completed our determination of the fixed point and the critical surface for $2<d<4$.

A plot of $t^*(\phi^2)$ vs ϕ^2 can be worked out numerically without difficulty. Figure 2 shows t^* for $d=3.8$, 3, 2.2. The probability distribution described by the fixed

[1] This assumption is not obviously necessary, but turns out to be important.

point is

$$P \propto \exp(-\mathcal{H}^*),$$

$$\mathcal{H}^* = \int d^dx[(\nabla\phi)^2 + U^*(\phi^2)], \tag{4.43}$$

$$U^*(\phi^2) = \int_0^{\phi^2} dx t^*(x). \tag{4.44}$$

We have plotted U^* vs ϕ^2 in Fig. 3.

For $d>4$, the conclusion will be quite different. The term ζs^{d-4} in (4.39) becomes large for large s. Note that the integral in (4.39) contains a term proportional to $t's^{d-4}$. Since $t' \approx u_c\zeta$ by (4.40), (4.39) becomes, for large s,

$$t's^{d-4} = O(1). \tag{4.45}$$

Thus, for $s\to\infty$, we have $t'\to t^*=0$. Thus, for $d>4$ the fixed point is the "trivial fixed point" and the corresponding $\mathcal{H}^*$ is simply

$$\mathcal{H}^* = \int d^dx(\nabla\phi)^2, \qquad d>4. \tag{4.46}$$

If d is exactly 4, all terms in (4.39) are of $O(1)$, plus a t' ln s term coming from the integral. Letting $s\to\infty$, the conclusion is still that

$$t'\to t^* \propto \lim_{s\to\infty} (\ln s)^{-1} = 0, \qquad d=4, \tag{4.47}$$

i.e., a trivial fixed point.

We shall not draw any conclusion for $d\leq 2$.

When $d<4$ but is very close to 4, t^* will be very close to zero. Equation (4.42) can be expanded in powers of t^* first and then t^* can be solved in powers of ϵ. Let us write t^* as a power series in ϕ^2:

$$t^*(\phi^2) = \sum_{m=0}^{\infty} u_{2(m+1)}(\phi^2)^m \tag{4.48}$$

[$u_{2(m+1)}$ here is actually $(m+1)$ times the $u_{2(m+1)}$ given by (4.2) and $u_2 = t_0$.]

One finds

$$u_2 = -(\epsilon\Lambda^2/2) + O(\epsilon^2)$$

$$u_4 = (16\pi^2\epsilon/n)\Lambda^\epsilon + O(\epsilon^2)$$

$$u_{2(m+1)} = (\epsilon\Lambda^2/2)[-(16\pi^2\epsilon/n)\Lambda^{-2+\epsilon}]^m \times (m-1)^{-1} + O(\epsilon^{m+2}), \qquad m>1. \tag{4.49}$$

We put the ϵ term in the power of Λ to remind the reader of the fact that ln Λ will appear if we expand in powers of ϵ even though it does not appear to the lowest order in ϵ [u_2, u_4 are of $O(\epsilon)$ while $u_6 = O(\epsilon^3)$, $u_8 = O(\epsilon^4)$, etc.]

E. Large s Behavior of R_s and Critical Exponents

Since we have an explicit formula for R_s for any s, the large s behavior can be extracted directly from (4.35), or, equivalently and more easily, from (4.40) and (4.39). There is no need to linearize and then look for eigenvalues.

Let t_1 be nonzero but very small and consider s to be large but not too large so that t' is still of $O(1)$. [Remember that $t_1=0$ means μ is on the critical surface. See (4.37).] With $t_1\neq 0$, (4.40) becomes

$$t' = s^2t_1 + u_c\zeta[1+O(\zeta/s^2)]$$

$$\therefore\ \zeta = -s^2t_1/u_c[1+O(t_1)] + O(s^{-2}). \tag{4.50}$$

Subtracting (4.42) from (4.39), we obtain

$$-(d-2)\Lambda^{2-d}\int_\Lambda^\infty dp p^{d-1}[(t'+p^2)^{-1} - (t^*+p^2)^{-1}] = (s^{d-2}t_1/u_c)(1+O(t_1)) + O(s^{d-4}). \tag{4.51}$$

If $t'-t^*$ is small compared to Λ^2, we obtain

$$t'-t^* \propto t_1 s^{d-2}(1+O(t_1)) + O(s^{d-4}), \tag{4.52}$$

which means, in the language of Sec. III [see (3.6)]

$$y_1 = d-2 \tag{4.53}$$

$$y_2 = d-4. \tag{4.54}$$

It can be easily shown that $y_3 = d-6$, $y_4 = d-8$, and so on. Clearly, only y_1 is positive and the rest are negative. The critical exponents follow:

$$\eta = 0$$

$$1/\nu \equiv y_1 = d-2,$$

$$y_2 = d-4. \tag{4.55}$$

These are true for $2<d<4$.

For $d>4$, we showed that $t^*=0$ [see (4.45)].

For small t', we substitute (4.50) in (4.39) to obtain

$$\frac{\phi^2}{N_c} = 1 + \left(-\frac{s^2t_1}{u_c} + O(s^{-2})\right)s^{d-4} + \frac{t'}{\Lambda^2}\left(\frac{d-2}{d-4}\right) \times (s^{d-4}-1). \tag{4.56}$$

Solving for t', we see that

$$t' \sim t_1s^2 + O(s^{4-d}), \qquad d>4, \tag{4.57}$$

$$\therefore\ \eta = 0,$$

$$1/\nu = 2,$$

$$y_2 = 4-d, \tag{4.58}$$

for $d>4$. For $d=4$, the last term of (4.56) would be proportional to ln s. It is left as an exercise.

F. Remarks

The above explicit calculation shows that the picture put forth in Sec. III is completely realized in the limit of large n for $d>2$. The most conspicuous feature of our results is the crucial dependence on the dimension d. The exponents η, y_1, y_2, $\cdots$ depend on nothing but d. (Of course, $n\to\infty$ here, that is why we do not see the n dependence.) In view of the arguments in Sec. III, we see that the universality of critical exponents is true for the large n limit.

In Sec. III, we introduced the notion of a critical region [see (3.12) and (3.16)], which is the region in k

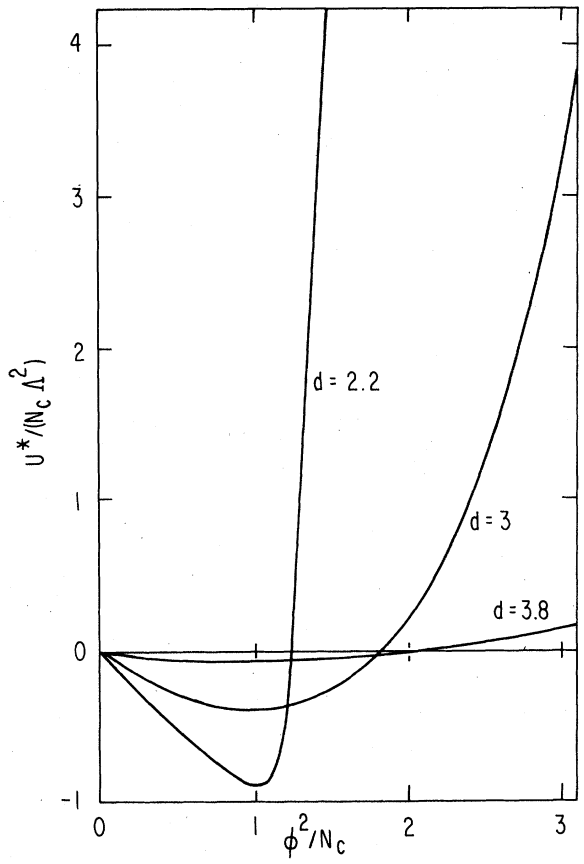

FIG. 3. $U^*(\phi^2)$ for $d=2.2$, 3, and 3.8. See (4.44).

space or the temperature range in the neighborhood of T_c in which the leading power of k or $T-T_c$ dominates the behavior of the correlation function or other singular quantities. The size of the critical region depends not only on the exponent y_2, but also on some of the details of $\mu(T)$. The criterion is based on how large an s is needed to get $R_s\mu(T)$ close to μ^*. The larger the s, the smaller the critical region.

To show that certain features of μ are more important than others in determining the size of the critical region, we work out more details of the $O(s^{d-4})$ term in (4.51). It takes some algebra to find for $t_1=0$ that

$$t'-t^*=\{u_c-[(4-d)/(d-2)](\Lambda^d/N_c)\}\times\text{constant } s^{d-4}+O(s^{d-6}), \quad (4.59)$$

where the "constant" depends on the fixed point, not on μ. The quantity u_c is defined by (4.41) and is a property of μ. Equation (4.59) shows that u_c is a crucial property. If u_c is adjusted so that the coefficient of s^{d-4} vanishes, then the critical region will be determined by the $O(s^{d-6})$ term, which is expected to be much smaller. Thus the critical region will be much larger.

V. GRAPH EXPANSION

Up to this point our discussion has been completely free of perturbation theory. As was emphasized in the Introduction, the renormalization group, in principle and in practice, is designed for nonperturbative analysis. This point must not be forgotten even though in some cases a perturbation expansion turns out to be helpful, and has been used extensively in the literature.

In the language of graphs, perturbation theory becomes a useful device. However, a trouble with a language is that it takes some time for one to become fluent even though one understands the rules immediately.

We shall first introduce the rules for graph expansion. Then the renormalization group defined in Sec. II will be expressed in graph language. In the next section, the renormalization group will be analyzed using graph language for the case of small ϵ.

A. Introducing Graphs

The integrations over a virtually infinite number of random variables ϕ_k are very difficult except when most of these random variables are statistically independent; i.e., when P is a product of distributions each involving only one or two random variables. The graph expansion starts by separating $\mathcal{H}$ into two pieces

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_I, \quad (5.1)$$

where

$$\mathcal{H}_0=\tfrac{1}{2}\sum_{k,i}|\phi_{ki}|^2 G_0^{-1}(k) \quad (5.2)$$

is the $m=1$ term in (2.16). We shall assume $u_2=\frac{1}{2}\delta_{i_1 i_2}G_0^{-1}$. The rest of $\mathcal{H}$ are included in $\mathcal{H}_I$. If $\mathcal{H}_I$ is ignored, $P\propto\exp(-\mathcal{H}_0)$ is a product of independent Gaussians since $\mathcal{H}_0$ is a sum of quadratic terms. Averages are easily computed. For example, we have $\langle\phi_{ik}\rangle_0=0$,

$$\langle\phi_{ik}\phi_{i-k}\rangle_0=\int d\phi_{ik}\,d\phi_{i-k}[\exp(-|\phi_{ik}|^2/G_0(k))]\times|\phi_{ik}|^2\left[\int d\phi_{ik}\,d\phi_{i-k}\exp(-|\phi_{ik}|^2/G_0(k))\right]^{-1},$$
$$=G_0(k), \quad (5.3)$$

where $\int d\phi_{ik}\,d\phi_{i-k}$ means integrating over the complex ϕ_{ik} plane. Note that $\phi_{ik}{}^*=\phi_{i-k}$ [see (2.38) and (2.39)]. We shall always denote such Gaussian averages by the subscript 0.

Now we write

$$\exp(-\mathcal{H})=\exp(-\mathcal{H}_0)\exp(-\mathcal{H}_1), \quad (5.4)$$

and any average $\langle A\rangle$ over the full distribution becomes

$$\langle A\rangle=\langle\exp(-\mathcal{H}_I)A_0\rangle/\langle\exp(-\mathcal{H}_I)\rangle_0=\sum_{n'=0}^{\infty}\frac{(-)^{n'}}{n'!}\langle\mathcal{H}_I{}^{n'}A\rangle_0\Big/\sum_{n'=0}^{\infty}\frac{(-)^{n'}}{n'!}\langle\mathcal{H}_I{}^{n'}\rangle_0. \quad (5.5)$$

Let us assume that A, as well as $\mathcal{H}_I$, are sums of products of the ϕ_k's. Since the Gaussian average of a

product of ϕ_k's is a product of pairwise averages [each ϕ_{ik} has to pair with a ϕ_{i-k} to give $\langle\phi_{ik}\phi_{i-k}\rangle_0 = G_0(k)$], the numerator and the denominator of (5.5) are complicated sums of products of $G_0(k)$'s. To introduce graphic representations, it is more convenient to use the random variables $\phi(x)$ [see (2.13)], instead of the ϕ_k's, because the coordinate space is easier to visualize. The Gaussian average of a product of $\phi(x)$'s (with different x's in general) is a sum of products of pairwise averages since $\phi(x)$ is a linear combination of ϕ_k's. Each pair gives, if we write $(2\pi)^{-d}\int d^dk$ for $L^{-d}\sum_{k<\Lambda}$,

$$\langle\phi_i(x)\phi_j(x')\rangle_0 = (2\pi)^{-d}\int d^dk G_0(k)\,\exp[ik\cdot(x-x')]\delta_{ij},$$
$$\equiv G_0(x-x')\delta_{ij}, \tag{5.6}$$

which can be represented by drawing a line between x and x'. Various averages can then be represented by graphs. As an illustration, suppose that

$$\langle A\rangle = \langle\phi_i(y)\phi_i(0)\rangle = G(y) \tag{5.7}$$
$$\mathcal{H}_1 = \tfrac{1}{2}u_4\int d^dx[\phi^2(x)]^2, \tag{5.8}$$

where ϕ^2 is given by (2.13), and $\mathcal{H}_1$ is just the last term of the Landau–Ginsburg form (2.12). Then (5.5) is a power series in u_4. To zeroth order in u_4, we simply have $G(y) = G_0(y)$. To first order, we have an additional term

$$u_4(\tfrac{1}{2}n+1)\int d^dx' G_0(y-x')G_0(x'-x')G_0(x') \tag{5.9}$$

as represented by Fig. 4(a). We use a dashed line for u_4 only to separate the two $\phi^2(x)$ factors in (5.8).[2] The second-order terms are given in Fig. 4(b). Those readers who are not familiar with graphs should write out the second-order terms explicitly. Note that disconnected graphs appear in both the numerator and denominator of (5.5). The net result is that only connected graphs contribute to $\langle A\rangle$. Note also that if A is of the form $A_1A_2\cdots A_m$ then there will be disconnected graphs of the form $\langle A_1A_2\cdots A_l\rangle\langle A_{l+1}\cdots A_m\rangle$ provided that neither of the two averages vanish. The coordinate representation is useful only for visualization. In practice, the wave vector representation is more convenient. Any random variable A to be averaged over is regarded as a product of ϕ_k's. So are powers of $\mathcal{H}_1$. Every line in a graph will be labeled by a wave vector. The sum over wave vectors is now a well-defined integral in k space. In each graph, those lines whose wave vectors are integrated over will be called *internal lines*. Those lines with wave vectors fixed by the ϕ_k's in A will be called *external lines*.

Of great importance is the "linked cluster theorem," which says that

$$\langle\exp(-\mathcal{H}_I)\rangle_0 = \exp\langle[\exp(-\mathcal{H}_I)-1]\rangle_{0c}, \tag{5.10}$$

[2] By this we mean the following. We imagine that the two factors of ϕ^2 are $\phi^2(x)\phi^2(x+\delta)$. The dashed line then joins x and $x+\delta$. The displacement δ is infinitesimal.

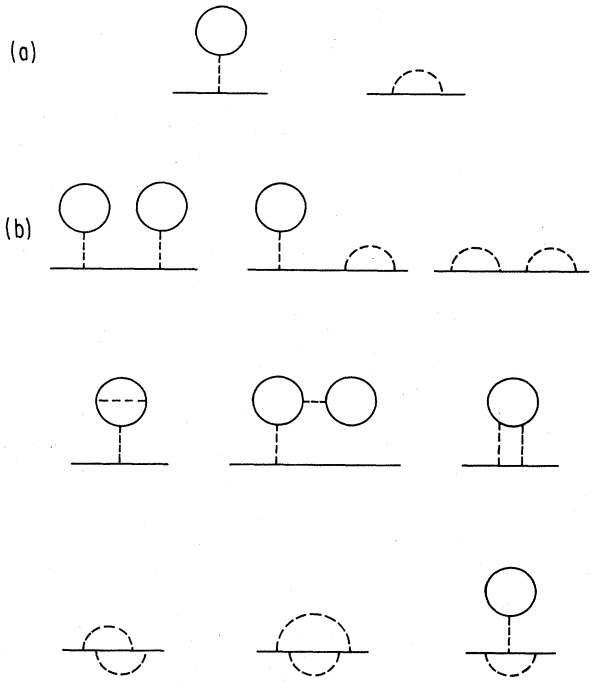

FIG. 4. Examples of graphs for G: (a) $O(u_4)$ terms, (b) $O(u_4^2)$ terms.

where the subscript c denotes the sum of connected graphs only. The disconnected graphs are generated by exponentiation. The proof is left as an exercise in counting graphs. We shall make some use of this theorem shortly.

A frequently occurring phrase is the "self-energy" Σ defined by

$$G^{-1}(k) = G_0^{-1}(k) + \Sigma(k). \tag{5.11}$$

The self-energy graphs are simply those graphs of $G(k)$ with G_0 lines of wave vector identical to k dropped.

What we have just gone through is the same as the Wick's theorem and Feynman graph expansion in field theory, if the time variable there is taken as imaginary and counted as a space dimension.

B. The Multiple Integral

The multiple integral in (2.21) is the first step in defining R_s. Let us denote those random variables to be integrated over by $\bar{\phi}$ and those not to be integrated by ϕ. To save writing, we shall introduce the notation

$$\int\delta\bar{\phi} \equiv \prod_{i,\Lambda/s<k'<\Lambda}\int d\phi_{ik'}. \tag{5.12}$$

We shall also write (2.16) as a sum:

$$\mathcal{H} \equiv \mathcal{H}(\phi) + \mathcal{H}(\phi,\bar{\phi}) \tag{5.13}$$

for the $\mathcal{H}$ in (2.21). Here $\mathcal{H}(\phi)$ is the part depending only on ϕ, and $\mathcal{H}(\phi,\bar{\phi})$ is the part depending on both ϕ and $\bar{\phi}$. More explicitly, $\mathcal{H}(\phi)$ is given by (2.16) with all

wave vectors restricted to less than Λ/s, and $\mathcal{H}(\phi, \bar{\phi})$ is the rest. The graphic representation of $\int \delta\bar{\phi} \exp(-\mathcal{H})$ can be introduced as we did previously. Similar to (5.1) and (5.2), we write

$$\mathcal{H}(\phi, \bar{\phi}) = \mathcal{H}_0(\bar{\phi}) + \mathcal{H}_I(\phi, \bar{\phi}), \tag{5.14}$$

$$\mathcal{H}_0(\bar{\phi}) \equiv \sum_{\Lambda/s<k<\Lambda} \tfrac{1}{2} \sum_{i=1}^{n} |\phi_{ik}|^2 G_0^{-1}(k). \tag{5.15}$$

We then define the Gaussian average by dropping $\mathcal{H}_I$ as before:

$$\langle A \rangle_{\bar{0}} \equiv \int \delta\bar{\phi} A \exp[-\mathcal{H}_0(\bar{\phi})] / \int \delta\bar{\phi} \exp[-\mathcal{H}_0(\bar{\phi})]. \tag{5.16}$$

The additional "bar" in the subscript of $\langle \cdots \rangle_0$ denotes that the average is taken over the random variables $\bar{\phi}$. Then the multiple integral in (2.21) can be written as

$$\int \delta\bar{\phi} \exp[-\mathcal{H}(\phi) - \mathcal{H}(\phi, \bar{\phi})] = \exp[-\mathcal{H}(\phi)] \langle \exp[-\mathcal{H}_I(\phi, \bar{\phi})] \rangle_{\bar{0}} \times \int \delta\bar{\phi} \exp[-\mathcal{H}_0(\bar{\phi})]. \tag{5.17}$$

The last factor is a constant independent of ϕ. The average in the middle of (5.17) can be expanded and represented by graphs:

$$\langle \exp[-\mathcal{H}_I(\phi, \bar{\phi})] \rangle_{\bar{0}} = \sum_{n=0}^{\infty} \frac{(-)^n}{n!} \langle \mathcal{H}_I{}^n(\phi, \bar{\phi}) \rangle_{\bar{0}} = \exp \langle \{\exp[-\mathcal{H}_I(\phi, \bar{\phi})] - 1\} \rangle_{\bar{0}c}, \tag{5.18}$$

where the last line follows from the linked cluster theorem (5.11). Remember that $\bar{\phi}$ denotes the random variables $\phi_{k'}$, with $\Lambda/s<k'<\Lambda$. Thus, the internal lines in the graphs now have wave vectors ranging between Λ/s and Λ in magnitude, i.e., wave vectors in a "shell" in k space. Now, if we substitute (5.18) and (5.17) in (2.21), we obtain $\mathcal{H}'$ apart from an additive constant

$$\mathcal{H}' = (\mathcal{H}(\phi) - \langle \{\exp[-\mathcal{H}_I(\phi, \bar{\phi})] - 1\} \rangle_{\bar{0}c})_{\phi_k \to \alpha_s \phi_{sk}}. \tag{5.19}$$

This is then the graphic representation of (2.21).

C. The Change of Scale; R_s Defined Graphically

The new parameters $\mu' = (G_0'^{-1}, u_4', u_6', \cdots)$ are now available in (5.19). For clarity, we shall extract μ' in two steps. First, let us write what is in the square bracket of (5.19) as

$$\mathcal{H}(\phi) - \langle \exp[-\mathcal{H}(\phi, \bar{\phi})] - 1 \rangle_{\bar{0}c} = \tfrac{1}{2} \sum_{i,k} |\phi_{ik}|^2 (G_0^{-1} + \Sigma_s) + \sum_{m=2}^{\infty} L^{-(m-1)d} \times \sum_{k_1 \cdots k_{2m-1}} \sum_{i_1 \cdots i_{2m}} \phi_{i,k_1} \cdots \phi_{i_{2m} k_{2m}} \bar{u}_{2m}, \tag{5.20}$$

i.e., we made an expansion in powers of unintegrated random variables. The wave vectors in (5.20) all have magnitude less than Λ/s. In terms of graphs, Σ_s is the self-energy; i.e., the sum of all graphs (connected, of course) with two external lines, and $\bar{u}_{2m}$ is the sum of all graphs of $2m$ external lines. *All internal lines of these graphs have wave vectors in the shell $\Lambda/s<k'<\Lambda$, while all external lines have wave vectors restricted to $k<\Lambda/s$.*

The second step is to replace ϕ_k by $\phi_{sk}\alpha_s$ and write sL' for L, $s^{1-\frac{1}{2}\eta}$ for α_s [see (2.30)] in (5.20). We obtain

$$\mathcal{H}' = \tfrac{1}{2} \sum_{i,k} |\phi_{ik}|^2 G_0'^{-1} + \sum_{m=2}^{\infty} L'^{-(m-1)d} \times \sum_{k_1 \cdots k_{2m-1}} \sum_{i_1 \cdots i_{2m}} \phi_{i,k_1} \cdots \phi_{i_{2m} k_{2m}} u_{2m}' \tag{5.21}$$

$$G_0'^{-1} = [G_0^{-1}(k/s) + \Sigma_s(k/s)] s^{2-\eta}, \tag{5.22}$$

$$u_{2m}' = \bar{u}_{2m} s^{-(m-1)d + m(2-\eta)}. \tag{5.23}$$

The quantity $\bar{u}_{2m}$ given by (5.20) of course depends on $k_1 \cdots k_{2m-1}$. In (5.23), it is understood that they are replaced by $k_1/s \cdots k_{2m-1}/s$, like the k in (5.22). Now in (5.21) the wave vectors range from 0 to Λ in magnitude, but, as mentioned before, the density of points in k-space is decreased from $L^d(2\pi)^{-d}$ to $L'^d(2\pi)^{-d}$. We now have a system of a smaller volume.

From $\mu = (G_0^{-1}, u_4, u_6, \cdots)$, which defines $\mathcal{H}$ via (2.16), we have arrived at (5.22) and (5.23) giving μ' by carrying out (2.21). We have thus established $\mu' = R_s \mu$ in terms of graphs.

D. The Exponent η and Self-Energy

In Sec. IIC we defined η with respect to a fixed point μ^*. We shall now observe a simple relationship between η and the derivative of the self-energy at the fixed point. Since $R_s \mu^* = \mu^*$, it follows from (5.22) that

$$G_0^{*-1}(k) = [G_0^{*-1}(k/s) + \Sigma_s^*(k/s)] s^{2-\eta}. \tag{5.24}$$

We expand $G_0^{*-1}(k)$ in powers of k^2:

$$G_0^{*-1}(k) = t_0^* + r_1^* k^2 + r_2^* k^4 + \cdots. \tag{5.25}$$

We can always choose the unit of k such that $r_1^* = 1$. Let us expand $\Sigma_s^*(k)$ also,

$$\Sigma_s^*(k) = \Sigma_s^*(0) + (\partial \Sigma_s^* / \partial k^2)_{k=0} k^2 + \cdots \tag{5.26}$$

and define

$$Z_s^{*-1} = 1 + (\partial \Sigma_s^* / \partial k^2)_{k=0}. \tag{5.27}$$

Then (5.24) reads

$$t_0^* + k^2 + \cdots = [t_0^* + \Sigma_s^*(0) + k^2 s^{-2} Z_s^{*-1} + \cdots] s^{2-\eta}. \tag{5.28}$$

Thus, we have

$$t_0^* = [t_0^* + \Sigma_s^*(0)] s^{2-\eta}, \tag{5.29}$$

$$Z_s^* = s^{-\eta}. \tag{5.30}$$

Therefore, $\eta = 0$ if $Z_s^* = 1$; i.e., if $\Sigma_s^*(k^2)$ is independent of k.

E. Technical Remarks

In this section we have assumed that $\langle\phi(x)\rangle=0$. If there is an external field or in the case where μ goes below the critical surface ($t_1<0$), this would no longer be true, and the graphs will have some additional features which can be easily included.

The smoothed cutoff mentioned in Sec. II can be done easily in terms of the graph language. There are many ways to do it. For example, in calculating Σ_s and $\bar{u}_{2m}$ of (5.20), we replace $G_0(k')$ for each internal line by

$$G_0(k')f(k'^2/\Lambda^2)[1-f(k'^2s^2/\Lambda^2)], \tag{5.31}$$

where

$$\begin{aligned} f(x)&=1-e^{-1/10(x-1)^2}, && x>1\\ &=1, && x<1 \end{aligned} \tag{5.32}$$

or any function dropping from 1 to zero over a narrow but finite region around $x=1$.

VI. THE RENORMALIZATION GROUP FOR THE CASE OF SMALL ϵ

In Sec. IV, where we worked out R_s for the case of large n, it was clear that the fixed point for small $\epsilon\equiv 4-d$ was of $O(\epsilon)$; i.e., very close to the trivial fixed point. This feature remains for arbitrary n, as was found by Wilson and Fisher (1972), who worked out the renormalization group for small ϵ. Since their discovery, a vast amount of literature has followed using the idea of expanding in powers of ϵ. An extensive treatment is included in Wilson and Kogut (1972).

In this section, we give the small-ϵ case as our second example of working out R_s explicitly. We shall examine the fixed point and its neighborhood. In deriving R_s, we shall make use of the graph expansion introduced in Sec. V.

A. The Parameter Space

We shall use the parameter space defined by the Landau–Ginzburg form (2.12), which is sufficient for the lowest-order terms in ϵ. Furthermore, the parameter a in (2.12) will be fixed at unity. This is because a will not change under R_s in the following discussion, similar to the case in Sec. IV. We therefore have, for the probability distribution,

$$P\propto\exp(-\mathcal{H}),$$

$$\mathcal{H}=\int d^dx((\nabla\phi)^2+t_0\phi^2+\tfrac{1}{2}u_4\phi^4), \tag{6.1}$$

where ϕ^2, ϕ^4 are defined by (2.13). Leaving out the parameter a in (2.15), we have a simple two-dimensional space of points

$$\mu=(t_0, u_4). \tag{6.2}$$

As will be clear later, this parameter space is sufficient for $O(\epsilon)$ terms only. If one wants to go beyond the first order in ϵ, the task will be much more difficult, and this simple parameter space will not suffice. One would have to include parameters for the wave vector-dependent terms of u_4, for example.

B. Formulas for R_s

We shall find $\mu'=R_s\mu$, regarding t_0 and u_4 as quantities of $O(\epsilon)$. We do not have to go back to (2.21) because now u_4 is small and the formulas (5.22) and (5.23) from the graph expansion become useful. All we need to do is to evaluate some lowest-order graphs.

To first order in ϵ, the two graphs in Fig. 4(a) are the only graphs for Σ_s. Since they are independent of the external wave vector, we must have

$$\eta=0 \tag{6.3}$$

in view of the conclusion below (5.30). Our G_0^{-1} is simply t_0+k^2. Thus, (5.22) says that

$$G_0'^{-1}\equiv t_0'+k^2=(t_0+\Sigma_s)s^2+k^2. \tag{6.4}$$

The two graphs for Σ_s give

$$\begin{aligned}\Sigma_s&=(\tfrac{1}{2}n+1)u_4\int[d^dp/(2\pi)^d](p^2+t_0)^{-1},\\ &=(\tfrac{1}{2}n+1)u_4K_{4-\epsilon}\int_{\Lambda/s}^{\Lambda}dpp^{3-\epsilon}(t_0+p^2)^{-1},\end{aligned} \tag{6.5}$$

where K_d is given by (4.23). Remember that $t_0=O(\epsilon)$. We can expand (6.5) in ϵ and substitute it in (6.4) to obtain

$$\begin{aligned}t_0'=s^2t_0+(\tfrac{1}{2}n+1)(u_4/8\pi^2)[\tfrac{1}{2}\Lambda^2(1-s^{-2})-t_0\ln s]s^2\\ +u_4\epsilon C+u_4^2D,\end{aligned} \tag{6.6}$$

where ϵC is a constant of $O(\epsilon)$ resulting from the $O(\epsilon)$ terms of $K_{4-\epsilon}$ and $p^{3-\epsilon}$; u_4^2D denotes second-order graphs which we have not included. The reason for writing the C and D terms in (6.6) is that they are of $O(\epsilon^2)$, the same magnitude as the term proportional to $u_4t_0\ln s$. While the C and D terms play no part in our final results, the $u_4t_0\ln s$ term does, as we shall see shortly.

Following (5.23), we construct u_4'. Besides u_4 itself, we must include in $\bar{u}_4$ all first-order corrections, which are shown in Fig. 5(a). If we ignore the dependence of $\bar{u}_4$ on the external wave vectors, all of these five graphs, except an additional factor $n/2$ for the graph with a closed loop, give the same contribution. This is easily seen by shrinking all the dashed lines to points. Then all five reduce to the form of Fig. 5(b). Before we write down (5.23) for u_4', we must observe the following. As was mentioned below (5.23), $\bar{u}_{2m}$ and u_{2m}' are functions of the wave vectors of external lines. We can expand both sides of (5.23) in powers of the external wave vectors. For each power, there will be an equation. Thus, (5.23) is really a set of equations for the coefficients of this power series expansion for $\bar{u}_{2m}$ and u_{2m}'. We shall keep track of only the first of these equations for u_4', namely, that for the constant term

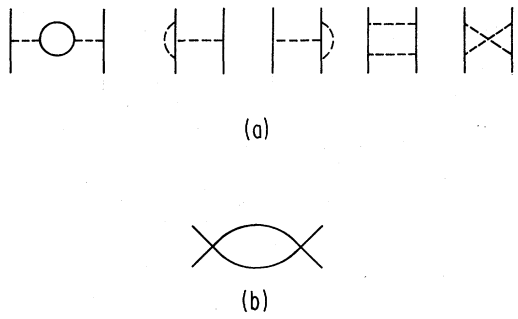

FIG. 5. Lowest-order corrections to u_4.

obtained by setting the external momenta to zero,

$$u_4' = \left(u_4 - (\tfrac{1}{2}n+4)u_4^2 \int_{\Lambda/s}^{\Lambda} \frac{d^d p}{(2\pi)^d} (t_0+p^2)^{-2}\right) s^\epsilon. \quad (6.7)$$

The rest of the set of equations will not concern us. To include them, our parameter space would have to be enlarged, as mentioned earlier. Since the u_4^2 term in (6.7) is of $O(\epsilon^2)$, we can set $d=4$ and $t_0=0$. The integral is trivial. We then have

$$u_4' = u_4[1-(\tfrac{1}{2}n+4)(u_4/8\pi^2)\ \ln s]s^\epsilon. \quad (6.8)$$

Our determination of $R_s\mu=(t_0', u_4')$ is thus complete: given (t_0, u_4), (6.6) and (6.8) furnish t_0' and u_4' to $O(\epsilon^2)$, apart from the unknown constants C and D in (6.6).

The reader may still question the consistency of the above procedure. We are excluding parameters of higher orders than $O(\epsilon)$; yet in (6.6) and (6.8) there are terms of $O(\epsilon^2)$. It is quite conceivable that we have missed a parameter of $O(\epsilon^2)$ which would modify (6.6) and (6.8). The major cause of worry would be u_6, which is expected to be of $O(\epsilon^2)$ and to modify (6.6) and (6.8) linearly. But luckily this is not the case. One piece of evidence is given by the large n result (4.49) which says that u_6^* is not of $O(\epsilon^2)$, but is of $O(\epsilon^3)$ instead. In fact, if we assume $u_6=0$ to start with, then R_s will be able to generate a u_6' of $O(\epsilon^2)$ which is non-zero only for very large external wave vectors over a very small range, and the inclusion of u_6 would not affect our results (6.6) and (6.8). Showing this in detail is left as an exercise, and is also given in Wilson and Kogut (1972). Another parameter of $O(\epsilon^2)$ which we are not keeping is the wave vector-dependent part of u_4. It will affect (6.6) and (6.8) only to $O(\epsilon^3)$.

C. The Fixed Point and Eigenvalues of the Linearized R_s

To find the fixed point, we set $\mu=\mu'=\mu^*=(t_0^*, u_4^*)$ in (6.6) and (6.8). Solving for u_4^* and t_0^*, we find

$$u^*/8\pi^2 = [2\epsilon/(n+8)]+O(\epsilon^2), \quad (6.9)$$

$$t_0^* = -(1+\tfrac{1}{2}n)(u_4^*/8\pi^2)\cdot\tfrac{1}{2}\Lambda^2+O(\epsilon^2),$$
$$= -[(n+2)/(n+8)]\tfrac{1}{2}\epsilon\Lambda^2+O(\epsilon^2). \quad (6.10)$$

The s^ϵ term in (6.8) can be replaced by $1+\epsilon \ln s$. The results (6.9) and (6.10) are consistent with the large n results (4.49). The C and D terms in (6.6) contribute only to the $O(\epsilon^2)$ terms in (6.10).

We proceed to linearize R_s and find the eigenvalues, etc. of R_s^L. Let $\delta t_0=t_0-t_0^*$, $\delta t_0'=t_0'-t_0^*$, etc. The linearized version of (6.8) is

$$\delta u_4' = \delta u_4[1-(\tfrac{1}{2}n+4)(2u^*_4/8\pi^2)\ \ln s]s^\epsilon$$
$$= \delta u_4 s^{-\epsilon}, \quad (6.11)$$

where we have used (6.9) for u^*_4. From (6.6), we obtain

$$\delta t_0' = \delta t_0[1-(u_4^*/8\pi^2)(\tfrac{1}{2}n+1)\ \ln s]s^2+\delta u_4 A,$$
$$= \delta t_0 s^{y_1}+\delta u_4 A, \quad (6.12)$$

where $y_1=2-(1+\frac{1}{2}n)u_4^*/8\pi^2$, and

$$A = (\Lambda^2/16\pi^2)(1+\tfrac{1}{2}n)(s^2-1)+O(\epsilon). \quad (6.13)$$

Equations (6.11) and (6.12) have the matrix form $\delta\mu'=R_s^L\delta\mu$ with

$$R_s^L = \begin{pmatrix} s^{y_1} & A \\ 0 & s^{y_2} \end{pmatrix}. \quad (6.14)$$

The eigenvalues are obviously s^{y_1} and s^{y_2}, with

$$1/\nu \equiv y_1 = 2-[(n+2)/(n+8)]\epsilon+O(\epsilon^2), \quad (6.15)$$

$$y_2 = -\epsilon. \quad (6.16)$$

The corresponding eigenvectors are easily found:

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} a \\ 1 \end{pmatrix} \quad (6.17)$$

$$a = -(1+\tfrac{1}{2}n)(\Lambda^2/16\pi^2)+O(\epsilon). \quad (6.18)$$

The critical surface here is thus the line passing through the fixed point in the e_2 direction. e_1 and e_2 are not orthogonal since R_s^L is not a symmetric matrix, and the terms C and D in (6.6) appear in the $O(\epsilon)$ term of A in (6.12) and (6.13). Although they contribute to (6.12) to the same order as the u_4^* term in the coefficient of δt_0, they have no effect on our answers for y_1, y_2 and e_1, e_2 to the order of interest.

The meaning of (6.18) which defines the critical surface may not be too clear. Let us go back to (6.6) and write it as

$$t_0'/s^2 = t_0+u_4(\tfrac{1}{2}n+1)(\Lambda^2/16\pi^2)(1-s^2). \quad (6.19)$$

The $O(\epsilon^2)$ terms are dropped. If $\mu=(t_0, u_4)$ is on the critical surface, then, for $s\to\infty$, $R_s\mu=\mu'$ must approach the fixed point (that was our definition of the critical surface), and we must have $t_0'/s^2\to t_0^*/s^2\to 0$. Thus, (6.19) becomes

$$0 = t_0+u_4(\tfrac{1}{2}n+1)(\Lambda^2/16\pi^2), \quad (6.20)$$

and defines the critical surface, a straight line whose slope is given by (6.18).

D. Remarks

The reader may feel unhappy about our replacing things like $1-\epsilon \ln s$ by $s^{-\epsilon}$. This would be quite wrong for large s. But isn't large s the very case of interest? Quite right. In fact, the error is very large if s is very large. Note that the form s^{y_i} for the eigenvalues of R_s^L is general; i.e., it follows from the fact that $R_s R_{s'} = R_{ss'}$. What we have been able to do is to find y_1, y_2 and the fixed point to $O(\epsilon)$. Of course, a small error in y_i will make a big error in s^{y_i} if s is very large.

Note also that $y_2 = -\epsilon$, which is very small if ϵ is very small; i.e., $O(s^{y_2})$ is not small unless s is exceedingly large. In the language of Sec. III [see (3.12), (3.16)], this means that the critical region is of the order $2^{-1/\epsilon}$ and is exceedingly small in general. This shows that, for small ϵ, universal critical behavior will be masked by nonuniversal model-dependent terms unless somehow the latter terms can be separated or removed. This point will come up again later.

Our result here is correct for any n. The reader is urged to compare it with those obtained in Sec. IV.

VII. THE BASIS FOR CALCULATION OF CRITICAL EXPONENTS BY PERTURBATION THEORY

The renormalization group analysis tells us how physical quantities such as the correlation function $G(k, \mu(T))$ should behave when T is very close to T_c, and how scaling laws appear. It tells us that the critical exponents appear as properties of R_s near its fixed point. We have seen explicitly how things work for the case of large n and also for the case of small ϵ. Hopefully, when ϵ becomes larger and n becomes smaller, the qualitative results will remain and the only changes would be quantitative changes in exponents and other numbers. When such a view is taken for granted, it becomes possible to compute the critical exponents by perturbation theory as power-series expansions either in $1/n$ or in ϵ, without studying the details of the renormalization group. Consider the following example. Since we know from renormalization group arguments that, at $T = T_c$, we have

$$G^{-1}(k) \propto k^{2-\eta}[1+O(k^{-y_2})], \tag{7.1}$$

and we have found that, for large n, $\eta = O(1/n)$ and $y_2 = d-4+O(1/n)$, then we can expand $k^{-\eta}$ in powers of $1/n$:

$$G^{-1}(k)k^{-2} \propto 1-\eta \ln k+\tfrac{1}{2}\eta^2 \ln^2 k+\cdots + O(k^{4-d}, n^{-1}k^{4-d} \ln k, \cdots). \tag{7.2}$$

Since this is true for *any* μ on the critical surface, we can pick the simplest one for the sake of computing η. We can pick

$$\mu_1 = (t_0+k^2, u_4, 0, 0, 0, \cdots), \tag{7.3}$$

namely, except u_2 and u_4, all other parameters are set to zero in (2.19). We shall choose $u_4 = O(1/n)$. Then we can compute $G(k)$ as a power series in $1/n$ by using a perturbation expansion of any kind. The quantity t_0 can be chosen to each order in such a way that $G(0)^{-1} = 0$; i.e., chosen to make sure that μ is on the critical surface to every order we calculate. The calculation of $G^{-1}(k)k^{-2}$ will result in a power series of $1/n \ln k$ and η is then identified by comparison with (7.2). The coefficients of the powers of $1/n \ln k$ will not depend on u_4, which appears only in the $O(k^{-y_2})$ term, as we argued in Sec. III. Recall that the $O(k^{-y_2})$ term reflects the approach to μ^* of $R_s\mu_1$ at a rate $s^{y_2} = s^{d-4+O(1/n)}$. This term is negligible in the critical region

$$k \ll 2^{-1/(4-d)}, \tag{7.4}$$

given by (3.12). As long as d is not close to 4, the size of the critical region is of $O(1)$. The above discussion is the basis for the $1/n$ expansion of critical exponents by perturbation theory, which has been studied extensively.

Similar arguments apply to the ϵ-expansion of critical exponents by perturbation theory, where one also starts with μ_1 with $u_4 = O(\epsilon)$. Troubles appear because the rate of approach of $R_s\mu_1$ to the fixed point μ^* is $s^{d-4} = s^{-\epsilon}$, for small ϵ; i.e., $y_2 = -\epsilon+O(\epsilon^2)$. The $O(k^{-y_2})$ term in (7.1), which has nothing to do with η, will also contribute a series in $\epsilon \ln k$, and one can no longer extract η by examining the coefficient of $\ln k$. In other words, the critical region vanishes in the small ϵ limit [see (IVD)]. One can get around this difficulty by choosing a special u_4 [$= u_0(\epsilon)$ of Wilson (1972)] so that the s^{y_2} term in $R_s\mu_1$ vanishes. Then $R_s\mu_1$ will approach μ^* at a rate s^{y_3} which is s^{d-6} in the large n limit and the small ϵ and any n. With this choice, the $O(k^{-y_2})$ term vanishes and the $O(k^{y_3}) = O(k^{2+O(\epsilon)})$ term will not give rise to any $\ln k$. Effectively, the critical region is extended to $O(1)$. This was how Wilson (1972) was able to compute critical exponents beyond the first order in ϵ by using perturbation theory. Note that the μ_1 so chosen is *not* the fixed point μ^*. It is a special point on the critical surface so chosen that $R_s\mu_1$ approaches μ^* at a rate s^{y_3} for large s instead of at the rate s^{y_2}. It has the nice feature that all entries but u_2 and u_4 are zero. Note also that no such special choice of u_4 is needed for the $1/n$ expansion where the $O(k^{y_2})$ term does not give rise to confusing logarithms. The u_4 will drop out automatically in the results for exponents.

VIII. SUMMARY AND DISCUSSION

A. Summary of Qualitative Points

Having gone through some quantitative calculation, we now make a brief summary of qualitative features. Let us think in terms of the spin-block picture of Kadanoff which we mentioned in the Introduction, namely, a lattice of blocks of spins. The size of a block is Λ^{-1}.

The interaction among the blocks of spins is described by a set of parameters μ. We emphasized in Secs. I and II that μ is a function of Λ. The renormalization group is a set of transformations R_s. Two steps are involved in R_s. First, we change Λ to Λ/s. This is a "coarse graining" procedure which increases the block size to $s\Lambda^{-1}$ by an averaging process. Second, we shrink the system by a factor s so that the block size returns to Λ^{-1}. In short, R_s is a change of scale with no change of the "spatial resolution," which is the block size Λ^{-1}. Under these operations, the set of parameters μ goes to a new set $\mu'=R_s\mu$. It is important to remember that the probability distributions described by μ and μ' are equivalent in the sense discussed in the beginning of Sec. II, and formulas such as (2.31)

$$G(k,\mu)=s^{2-\eta}G(k,\mu') \tag{8.1}$$

for the correlation function follow as a result of this equivalence.

Clearly, carrying out R_s simply changes one set of parameters to another; i.e., replaces the old problem by a new one. It does not *solve* the problem at all. The new one is just as difficult to solve as the old one. There will be applications where the new problem does become simpler. However, the spirit here is that we try to gain through R_s without solving the problem. It is illustrated in Sec. III that much about critical phenomena can be said through a formula like (8.1) if R_s has a fixed point μ^*. If $\mu(T)$ is the set of parameters given by the canonical ensemble for the system at temperature T, then $R_s\mu(T_c)$ will approach μ^* for large s. The fixed point μ^* describes a spin-block system which will not change under a scale change and coarse graining. Thus, a physical picture of the system at T_c follows. If we look at a system at T_c through a microscope with sufficiently poor spatial resolution, it would *look the same* if the magnification of the microscope were decreased. Another conclusion from Sec. III is that, if T is slightly above T_c, there is only one relevant parameter ξ, the correlation length, describing how far μ is from μ^* in a special direction e_1 in the parameter space. For large s while not too far away from μ^*, we have

$$R_s\mu(T)\sim(s/\xi)^{1/\nu}e_1+O(s^{y_2}), \tag{8.2}$$

with $\xi\propto|T-T_c|^{-\nu}$. Shrinking the system (increasing s) thus effectively *decreases* the correlation length. The scaling hypothesis follows when $O(s^{y_2})$ is neglected.

Our analyses of the large n case and the small ϵ case are just illustrations of these qualitative features. We emphasize that these are the simplest illustrations, in which we have assumed short-range interaction and maximum symmetry, in addition to the restriction of large n or small ϵ.

B. Concluding Remarks

The formalism presented in this article is by no means the unique formulation of the renormalization-group idea. There is a great deal of flexibility in formulation. The idea should be formulated according to the special problem involved and according to what kind of approximations one has in mind. The one given here is just one possibility.

It should be emphasized again that, in dealing with R_s, we transform the coupling parameters and study the patterns in the new parameters. The traditional approach to statistical mechanical models of critical phenomena is to solve them exactly with various ingenious mathematical tricks. Without these tricks, one faces serious difficulties coming from various singularities which are inherent in critical phenomena owing to large scale fluctuations. The situation is that one finds either the exact answer or no answer at all. There are few ways to make approximations (apart from numerical calculations and series expansions). In the renormalization-group approach, one does not try to solve the model. Instead, one deals with the coupling parameters, which are nonsingular. The singularities associated with the critical point are now reflected through large s in formulas like (8.1). Since no singularities are involved in R_s, we can carry out the calculation for R_s approximately in more simple-minded ways. This is a distinctive and valuable feature of the renormalization-group approach.

Finally, we want to emphasize that the renormalization group is a general scheme for studying large systems, not something designed just to calculate critical exponents. The field is wide open.

C. Some Confusion of Terminologies

It might be helpful to clear up some confusion in terminologies used by relativistic-field theorists and those used by statistical mechanics and solid-state physicists.

The term "to renormalize," as originated in field theory, means "to normalize again" not "to include interactions." Thus, for example, the exact correlation function $G(k,\mu)$ should be called "unrenormalized." The *renormalized* one is defined as

$$G_\lambda(k,\mu)\equiv[G(k,\mu)/G(\lambda,\mu)]A_\lambda, \tag{8.3}$$

which is normalized to A_λ at an arbitrarily chosen reference point λ. The parameter A_λ is also arbitrarily chosen. More renormalized quantities and parameters can be introduced for other kinds of correlation functions. For a "renormalizable" theory, the original parameters μ can be eliminated for $\Lambda\to\infty$ and only the renormalized quantities and parameters like A_λ remain. Eventually, these parameters will be fixed by experimental data, which would then fix the physical content of the theory.

The conventional renormalization group in relativistic field theories is defined as the transformations of the parameters such as A_λ under a change of λ (keeping the physical content fixed). Thus, A_λ plays the role of μ

and λ plays the role of Λ in our discussion in this article. This is a point where some confusion would result if not clarified.

It has become a common convention for many in solid-state theory and statistical mechanics that the word "renormalized" means "with all effects of interaction included (without further normalization)." We feel that the word "exact" is *shorter* and more appropriate than the words "renormalized" or "unrenormalized."

The term "renormalization" would be better saved for discussing changes of parameters in a theory under a change of Λ, or λ, or other references.

D. A Brief Guide to Some Recent Work

The scope of this article is very limited. Many important areas have not been touched. An extensive list of references can be found in Wilson and Kogut (1972). In the following, we shall mention a few areas where much work has been done recently, and where the renormalization-group idea plays a role. This is not meant to be a complete survey. Many of the references are not published yet, and the ones we happen to know may not be representative. Our purpose here is only to give the reader some rough idea of the present status.

A. Numerical Investigation:
Wilson (1971),
Grover, Kadanoff, and Wegner (1972),
Golner (1972, 1973).
We feel that this is a very important area where much remains unexplored.

B. Long-Range Force Interaction:
Fisher, Ma, and Nickel (1972),
Sak (1973),
Suzuki (1972).
Dipole interaction:
Aharony (1973),
Aharony and Fisher (1973),
Fisher and Aharony (1973).

C. ϵ Expansion of Critical Exponents:
Wilson and Fisher (1972),
Wilson (1972),
Wilson and Kogut (1972),
Nickel (1972),
Fisher and Pfenty (1972),
Brezin, Le Guillou, Zinn-Justin, and Nickel (1973).
Equation of State, and below T_c:
Brezin, Wallace, and Wilson (1973).

D. $1/n$ Expansion of Critical Exponents:
Abe (1972, 1973),
Abe and Hekami (1973),
Wilson (1972),
Ma (1972), (1973a),
Ferrell and Scalapino (1972),
Suzuki (1972, 1973).
Equation of State, and below T_c:
Brezin and Wallace (1973).

E. Correction to Scaling Laws, General Discussion:
Wegner (1972, 1973).

F. Application to Tricritical Phenomena:
Riedel (1972),
Riedel and Wegner (1972),
Wegner and Riedel (1973).

G. Nonspherical Symmetry:
Wallace (1973),
Aharony (1973).

H. Time-Dependent Phenomena at Critical Point:
Halperin, Hohenberg, and Ma (1972),
Suzuki (1973b).

I. There is a large literature on the conventional renormalization group in field theory, and more recently the approach via the Callan–Symanzik equation. See references cited in Wilson and Kogut (1972), and also the work of Brezin, Le Guillou, and Zinn-Justin (1973) and Di Castro *et al.* (1973).

A model for which Wilson's recursion formula becomes analytically tractable was discussed by Baker (1972).

Soda (1970) discussed the realization of a renormalization group by summing parquet diagrams. There is a rather large literature on the graph summing approach to critical phenomena. See Tsuneto and Abrahams (1973) and references cited therein.

ACKNOWLEDGMENTS

I thank Professor K. A. Brueckner, Professor D. Pines, Dr. J. J. Rehr, Mr. L. Senbetu, Professor J. C. Wheeler, and colleagues at UCSD for their helpful comments. I also thank Mr. Senbetu for his careful reading of the final draft.

*Alfred P. Sloan Foundation Fellow.
†Research supported in part by the National Science Foundation under Grant GP-38627X.

Abe, R., 1972, Prog. Theor. Phys. **48**, 1414.
Abe, R., 1973, Prog. Theor. Phys. **49** (1), ,
Abe, R., 1973a, preprints, Tokyo University (1973).
Abe, R., and S. Hikami, 1973, Phys. Lett. A **42**, 419.
Abe R., and S. Hikami, 1973a, Prog. Theor. Phys. **49** (2), ,
Aharony, A., 1973, preprints, Cornell University (1973).
Aharony, A., and M. E. Fisher, preprint Cornell University (1973).
Baker, G. A., Jr., 1972, Phys. Rev. B **5**, 2622.
Berlin, T. H., and M. Kac, 1952, Phys. Rev. **86**, 821.
Bogoliubov, N. N., and D. V. Shirkov, 1959, *Introduction to the Theory of Quantized Fields* (Interscience, New York 1959).
Brezin, E., J. C. Le Guillou and J. Zinn-Justin, 1973, preprints, CEN Saclay, France (1973).
Brezin, E., and D. J. Wallace, 1973, Phys. Rev. B **7**, 1969.
Brezin, E., D. J. Wallace, and K. G. Wilson, 1973, Phys. Rev. B **7**, 232.
Di Castro, C., G. Jona-lasinio, and L., Peliti, 1973, preprint,

Instituto di Fisica dell'Universitá, Padova, Italy (1973).
Ferrell, R. A., and D. J. Scalapino, 1972, Phys. Rev. Lett. **29,** 413.
Fisher, M. E., 1967, Rep. Prog. Phys. **30,** 615.
Fisher, M. E., and A. Aharony, 1973, Phys. Rev. Lett. **30,** 559.
Fisher, M. E., S. Ma, and B. G. Nickel, 1972, Phys. Rev. Lett. **29,** 917.
FIsher, M. E., and P. Pfenty, 1972, Phys. Rev. B **6,** 1889.
Gell-mann, M., and F. E. Low, 1954, Phys. Rev. **95,** 1300.
Golner, G. R., 1972, Ph.D. thesis, Cornell University (1972).
Golner, G. R., 1973, Phys, Rev. B, July (1973).
Golner, G. R., 1973a, preprint, Brookhaven National Laboratory, BNL-17681 (1973).
Grover, M. K., 1972, Phys. Rev. B **6,** 3546.
Grover, M. K., 1973, preprints, Brown University (1973).
Grover, M. K., L. P. Kadanoff, and F. J. Wegner, 1972, Phys. Rev. B **6,** 311.
Halperin, B. I., P. C. Hohenberg, and S. Ma, 1972, Phys. Rev. Lett. **29,** 1548.
Hikami, S., 1973, preprint, Tokoyo University (1973).
Kadanoff, L. P., 1966, Physics (N.Y.) **2,** 263.
Kadanoff, L. P., W. Götze, D. Hamblen, R. Hecht, E. A. S. Lewis, V. v. Palciauskas, M. Rayl, J. Swift, D. Aspnes, and J. Kane, 1967, Rev. Mod. Phys. **39,** 395.
Ma, S., 1972, Phys. Rev. Lett. **29,** 1311.
Ma, S., 1973a, Phys. Rev. A7, 2172 (1973).
Ma, S., 1973b, Phys. Lett. A **43,** 475.
Ma, S., 1973c, preprint, U. C. San DIego (1973).
Nickel, B. G., 1972, preprint, Cornell University (1972).
Riedel, E., 1972, Phys. Rev. Lett. **28,** 675.
Riedel, E., 1973, preprint, Duke University (1973).
Riedel, E., and F. J. Wegner, 1972, Phys. Rev. Lett. **29,** 349.
Sak, J., 1973, preprint, Cornell University (1973).
Soda, T., 1970, preprint, Tokyo University of Education (unpublished) (1970).
Stanley, H. E., 1968, Phys. Rev. **176,** 718.
Stanley, H. E., 1971, *Introduction to Phase Transitioons and Critical Phenomena* (Oxford U. P., New York, 1971).
Suzuki, M., 1972, Phys. Lett. A **42,** 5.
Suzuki, M., 1973a, Prog. Theor. Phys. **49,** 424. Prog. Theor. Phys. **49,** 1017. ; Prog. Theor. Phys. **49,** 1106.
Suzuki, M., 1973b, Phys. Lett. A **43,** 245.
Tsuneto, T., and E. Abrahams, 1973, Phys. Rev. Lett. **30,** 217.
Wallace, D. J., 1973, preprint, University of Southampton (1973).
Wegner, F. J., 1972a, Phys. Rev. B **5,** 4529.
Wegner, F. J., 1972b, Phys. Rev. B **6,** 1891.
Wegner, F. J., and E. Riedel, 1973, Phys. Rev. B **7,** 248.
Wilson, K. G., 1971, Phys. Rev. B **4,** 3174. Phys. Rev. B **4,** 3184.
Wilson, K. G., 1972a, Phys. Rev. Lett. **28,** 548.
Wilson, K. G., 1972b, preprint, CLNS-198, Cornell University (1972).
Wilson, K. G., and M. E. Fisher, 1972, Phys. Rev. Lett. **28,** 240.
Wilson, K. G., and J. Kogut, 1972, preprint, Insttute for Advanced Study, to be published in Phys. Report (1972).