3
The renormalization group idea

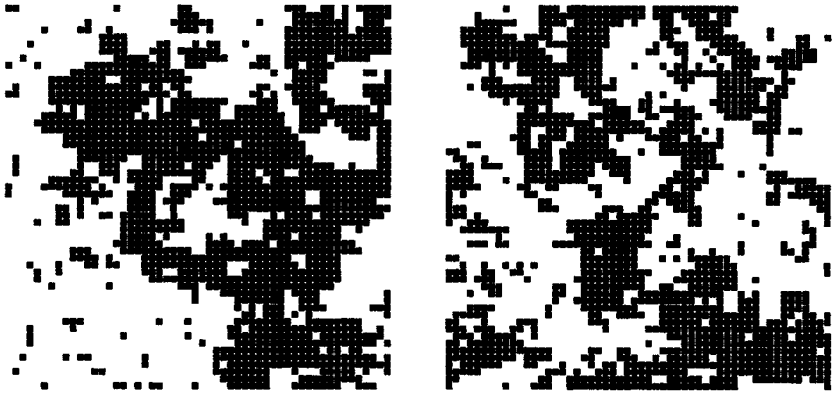
In this chapter the basic concepts of the modern approach to equilibrium critical behaviour, conventionally grouped under the title 'renormalization group', are introduced. This terminology is rather unfortunate. The mathematical structure of the procedure, in the sense that it may be said to have any rigorous underpinnings, is certainly not that of a group. Neither is renormalization in quantum field theory an essential element, although it has an intimate connection with some formulations of the renormalization group. In fact, the renormalization group framework may be applied to problems quite unrelated to field theory. The origins of the name may be traced to the particle physics of the 1960s, when it was optimistically hoped that everything in fundamental physics might be explained in terms of symmetries and group theory, rather than dynamics. One of the earliest applications of renormalization group ideas, in fact, was to the rather esoteric subject of the high energy behaviour of renormalized quantum electrodynamics. It took the vision of K. Wilson to realise that these methods had a far wider field of application in the scaling theory of critical phenomena that was being formulated by Fisher, Kadanoff and others in the latter part of the decade. By then, however, the name had become firmly attached to the subject.

Not only are the words 'renormalization' and 'group' examples of unfortunate terminology, the use of the definite article 'the' which usually precedes them is even more confusing. It creates the misleading impression that the renormalization group is a kind of universal machine through which any problem may be processed, producing neat tables of critical exponents at the other end. This is quite false. It cannot be stressed too strongly that the renormalization group is merely a framework, a set of ideas, which has to be adapted to the nature of the problem at hand. In particular,

whether or not a renormalization group approach is quantitatively successful depends to a large extent on the nature of the problem, but lack of such success does not necessarily invalidate the qualitative picture it provides.

All renormalization group studies have in common the idea of re-expressing the parameters which define a problem in terms of some other, perhaps simpler, set, while keeping unchanged those physical aspects of the problem which are of interest. This may happen through some kind of coarse-graining of the short-distance degrees of freedom, as in the problem of critical phenomena, where the long-distance physics is of interest. It may represent some kind of modification of the effects of large-scale disturbances, as in fluid turbulence, where the emphasis is on how such fluctuations are fed down to smaller distance scales. Or, in time-dependent problems, such as the dynamics of phase ordering following a quench from a disordered phase, it may correspond to the temporal evolution of the parameters specifying the early-time history, in such a way that the late time properties are left unaltered.

Whatever the motivation, these methods all end up with mathematical equations describing *renormalization group flows* in some complicated parameter space. It is the study of these flows, and what they tell us about the physical problem, which is the essence of renormalization group theory. In the context of equilibrium critical behaviour, this general aspect of the renormalization group appears most directly in the method of *real space* renormalization as applied to lattice spin systems, and it is with this example of the renormalization group in action that we shall therefore begin. It turns out that these real space methods are difficult to control in a quantitative fashion, as there is really no small parameter in which to expand. However, this feature does not weaken the remarkably powerful implications for scaling and universality which arise as quite general properties, and it is these consequences we wish to stress. Later, in Chapter 5, we shall describe a slightly different type of renormalization group, in which a small parameter, related to the number of dimensions of space, does appear, and which may therefore be used to yield systematically improvable quantitative results.



(a) (b)

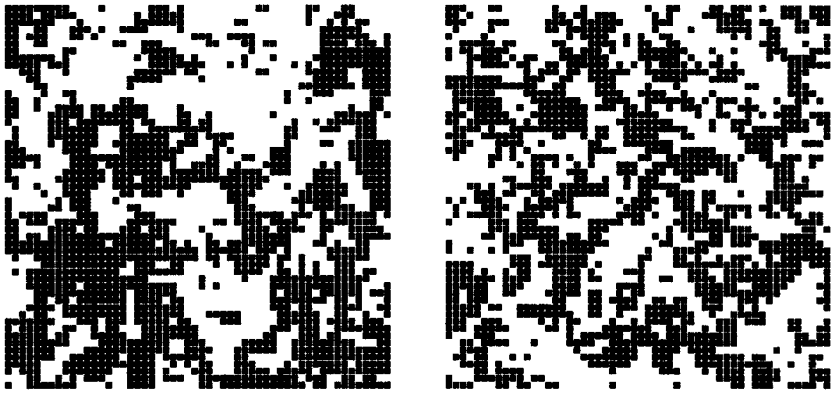
Figure 3.1. Typical configuration of the Ising model at $T = T_c$ (a), and (b) the result of a single block spin transformation.

3.1 Block spin transformations

Let us study the accompanying snapshots (Figures 3.1, 3.2) of typical states of the two-dimensional Ising model in zero magnetic field. They were produced by computer simulation. The first picture was taken close to the critical point, and we see clusters of down spins ($s = -1$) of all sizes. This is what makes the analysis difficult.

Now suppose we put the picture slightly out of focus, so that we can no longer see very well the microscopic details.† A mathematical way to implement this defocussing, or *coarse-graining*, is to make a *block spin transformation*. For definiteness, group the squares into 3×3 blocks, each containing 9 spins. To each block assign a new variable $s' = \pm 1$ whose value will indicate whether the spins in the block are predominantly up or down. The simplest

† In a lecture theatre this is rather easy to simulate with an overhead projector. Hyperopic readers should also have no difficulty in performing the experiment.



(a) (b)

Figure 3.2. Same as Figure 3.1, slightly above the critical temperature.

method is to take the 'majority rule' whereby $s' = +1$ if there are more spins up than down, and *vice versa*. When this is done, and the whole picture is rescaled by a linear factor of 3 so that the blocks are the same size as the original squares, we get the second picture. (It is the same overall size as the first picture because not all the system is shown there. In fact Figure 3.1a corresponds to the top left hand corner of Figure 3.1b.) The first thing to be noticed is that the second picture looks very much like the first. In fact, they are *statistically* the same, in that Figure 3.1b is an equally probable configuration for the critical Ising model as is Figure 3.1a. If we continue this blocking procedure, and start exactly at the critical point with a big enough system, all the pictures look pretty much the same. This observation illustrates the *scale invariance* of the critical system. On the other hand, if we start slightly above the critical temperature (Figure 3.2), although the original system may look very similar to that in Figure 3.1a, after a few transformations it soon looks very different.

All this is very qualitative, but it is nevertheless the essential basis of the renormalization group approach. If we were to start from a large sample of typical configurations, we could, by averaging over this sample, calculate all the correlation functions which characterise the system. These are, of course, determined in principle by the hamiltonian $\mathcal{H}(s)$. In the blocked system we could also measure the correlation functions of the block spins s'. In their turn, these correlation functions may be thought of as determined by some new hamiltonian $\mathcal{H}'(s')$. We can always cook up some arbitrarily complicated $\mathcal{H}'$ which will do this, but of course there is no reason to assume that it will be given simply as a sum over nearest neighbour exchange interactions. In general, it will include interactions between arbitrarily distant block spins s'. A basic assumption of the renormalization group, however, is that, no matter how many times the blocking transformation is iterated, the *dominant* interactions will be short-ranged. (Later, in Section 4.3, we shall consider the effect of longer ranged interactions and see how this statement can be made more quantitative.) Rather than attempting to prove the validity of this central assumption for each system of interest, its correctness is best borne out by its consequences of scaling and universality, which have been tested in many real experiments and numerical simulations, as well as being verified for those models which are exactly solvable.

Let us define the block hamiltonian $\mathcal{H}'$ more explicitly. The original system is described by a partition function

$$Z = \mathrm{Tr}_s e^{-\beta\mathcal{H}(s)}. \tag{3.1}$$

In what follows, we shall always absorb the factor of $\beta = 1/k_B T$ into the definitions of the various parameters in $\mathcal{H}$, such as the exchange coupling J and the magnetic field H. This defines what is called the *reduced* hamiltonian. Our majority rule may be implemented by inserting a projection operator under the trace, as follows. Define, for each block,

$$T(s'; s_1, \ldots, s_9) = \begin{cases} 1, & \text{if } s' \sum_i s_i > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

The new hamiltonian is then defined by

$$e^{-\mathcal{H}'(s')} \equiv \mathrm{Tr}_s \prod_{\text{blocks}} T(s'; s_i)\, e^{-\mathcal{H}(s)}. \tag{3.3}$$

Note that, in particular, because $\sum_{s'} T(s'; s_i) = 1$,

$$\mathrm{Tr}_{s'} e^{-\mathcal{H}'(s')} = \mathrm{Tr}_s e^{-\mathcal{H}(s)}, \tag{3.4}$$

that is, the partition functions Z for the original system and the blocked system are the same. But the above transformation preserves far more than this. Equation 3.3 implies that the whole probability distribution of quantities which depend only on spins $s', s'', s''', \ldots$ at higher levels of blocking will be left invariant. These include all the long wavelength degrees of freedom. Thus the whole of the large distance physics of the problem is left untouched by the renormalization group procedure. The only difference is that it should be expressed in terms of blocked, or renormalized, spins, rather than the original, or bare, spins.

It is useful to think of the couplings in the reduced hamiltonian $\mathcal{H}$ as forming a vector $\{K\} \equiv (K_1, K_2, \ldots)$. In the original model there might have been only one nearest neighbour coupling, say K_1, with all the other $K_i = 0$. But, as discussed above, the renormalization group will in principle generate all other possibilities. We may therefore picture the renormalization group transformation as acting on the space of all possible couplings $\{K\}$:

$$\{K'\} = \mathcal{R}\{K\}. \tag{3.5}$$

In the case of the Ising model, we may divide this very large space into the subspace of even couplings, which multiply interaction terms in the hamiltonian which are invariant under $s(r) \to -s(r)$, and the space of odd couplings, such as an external magnetic field.† If no odd couplings are present in the original model, none should be generated under renormalization.

All of this is rather general. Unfortunately, the sums involved in actually carrying out the trace over the s in (3.3) are intractable, and we must rely on some approximation scheme to proceed further with these block spin methods. It is not the purpose of this book to discuss such schemes, since the approximations involved are difficult to control. Nevertheless, the *qualitative* picture of the renormalization group which we extract from these considerations applies independently of any approximation. It is useful, all the

† In the case of a more general symmetry, each subspace corresponds to those interactions which transform according to an irreducible representation of the appropriate symmetry.

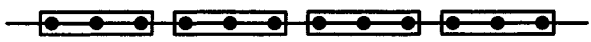


Figure 3.3. Blocking transformation for the one-dimensional Ising model.

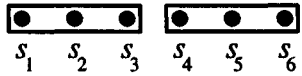


Figure 3.4. Two neighbouring blocks in the one-dimensional Ising model.

same, to examine at least one case in which they may be carried through exactly.

3.2 One-dimensional Ising model

In one dimension, the block spin renormalization group described above may often be carried through explicitly. Consider, for example, the simple zero-field Ising spin chain, with a reduced hamiltonian $\mathcal{H} = -K \sum_i s_i s_{i+1}$. This model may, of course, be solved by other means, for example, by using a transfer matrix, or, even more simply, by defining new variables $\sigma_i \equiv s_i s_{i+1}$. However, for the purposes of illustration, let us consider grouping the sites into blocks, each containing 3 spins, as shown in Figure 3.3. We could use the same majority rule as before, but it is analytically even simpler if we do something rather undemocratic, and count the vote of only the central spin in each block. This corresponds to taking $T(s'; s_1, s_2, s_3) = \delta_{s', s_2}$. The reason that this is justified is that, at very low temperatures, where all the action takes place, all the spins in a given block tend to vote the same way anyway. Thus this renormalization group transformation corresponds simply to performing the trace over the spins at the ends of each block, and leaving the central ones untouched. Such a procedure is called *decimation*. It works very well in one dimension.† Consider two neighbouring blocks, shown in Figure 3.4. Suppose we sum over the spins s_3 and s_4, keeping $s_1' \equiv s_2$ and $s_2' \equiv s_5$ fixed. The

† Since in this method the block spins are a subset of the original spins, it leads to the paradoxical result that the spin correlation function is constant at the critical point, which is incorrect in higher dimensions.

factors in the partition sum involving these degrees of freedom are

$$e^{Ks'_1 s_3}\ e^{Ks_3\ s_4}\ e^{Ks_4\ s'_2}. \tag{3.6}$$

Write $e^{Ks_3 s_4} = \cosh K(1+xs_3s_4)$, where $x \equiv \tanh K$, and similarly for the other two factors, giving

$$(\cosh K)^3(1 + xs'_1s_3)(1 + xs_3s_4)(1 + xs_4s'_2). \tag{3.7}$$

Now imagine expanding out this expression. When we perform the sum over $s_3, s_4 = \pm 1$, only terms with an even power of these variables will survive. We are therefore left with only

$$2^2(\cosh K)^3(1 + x^3s'_1s'_2), \tag{3.8}$$

which, apart from the constant outside, has the form of a nearest neighbour Boltzmann factor $e^{K's'_1s'_2}$, with

$$K' = \tanh^{-1}\left[(\tanh K)^3\right]. \tag{3.9}$$

The partition function of the whole system may thus be written in the required form $Z = \mathrm{Tr}_{s'} e^{-\mathcal{H}'(s')}$, where

$$\mathcal{H}'(s') = Ng(K) - K'\sum_i s'_i s'_{i+1}, \tag{3.10}$$

where N is the total number of original sites, and

$$g(K) = -\tfrac{1}{3}\ln\left[\frac{(\cosh K)^3}{\cosh K'}\right] - \tfrac{2}{3}\ln 2. \tag{3.11}$$

The renormalized hamiltonian, has, in this case, the same form as the original one, with a renormalized value of the coupling K', apart from a term independent of the s'_i. This additional term proportional to $g(K)$ does not affect the calculation of any expectation values, but it will enter into a calculation of the total free energy. It represents the contribution to the free energy from the short wavelength degrees of freedom which have been traced out. As we shall see in Sections 3.4 and 3.9, it plays an important role in the further development of the theory.

The content of the renormalization group transformation is expressed by the *renormalization group equation* (3.9). This is much more easily analysed in terms of the variable $x = \tanh K$, for which the renormalization group equation is simply $x' = x^3$. Recalling that K contains a factor $1/k_BT$, we see that high temperatures correspond to $x \to 0+$, and low temperatures to $x \to 1-$. Now suppose we iterate the process. Unless we begin with x exactly

Figure 3.5. RG flow for the one-dimensional Ising model.

equal to 1 ($T = 0$), it ultimately approaches zero. This means that the long distance degrees of freedom are described by a hamiltonian where the effective temperature is high, and we expect such a system to be in a paramagnetic state, with a finite correlation length. Since any system with $x < 1$ ultimately renormalizes into this region, we conclude that this whole region is paramagnetic. Only exactly at zero temperature is this not true. In terms of the renormalization group equation, we may say that there are two *fixed points*: the one at $T = 0$ is unstable, since any perturbation away from this is amplified by the renormalization group; the fixed point at $T = \infty$ ($x = 0$) is stable, and is the attractive fixed point for the whole region $0 \leq x < 1$. Every point in this region is therefore in the same phase, which is paramagnetic.† The *renormalization group flows* go from the unstable fixed point to the stable one, as shown in Figure 3.5.

This reflects a well-known fact about the one-dimensional systems with short-range interactions: they cannot be in an ordered state for $T > 0$, and therefore can have no true phase transition. This tendency towards the disappearance of order as the number of dimensions is lowered is characteristic of all systems. The dimension d_l such that systems in a given universality class have no phase transition for $d \leq d_l$ is called the *lower critical dimension.* For systems like the Ising model with discrete symmetries, $d_l = 1$. As we shall see in Section 6.1, for continuous symmetries $d_l = 2$ in general. This behaviour makes the above example rather uninteresting for our purposes, but, on the other hand, the analysis was rather simple. To illustrate how the renormalization group can be useful even when there is no true critical behaviour, however, let us use the above results to compute the correlation length ξ.

† Notice that we needed to supply additional physical input to describe the *nature* of the phase, which is not determined solely by the renormalization group.

The correlation length has, of course, the dimensions of length, but to express it as a pure number we may measure it in units of the lattice spacing a. In these units, it may depend on only the reduced coupling K. After performing the renormalization group transformation, the long distance physics is preserved, and so the dimensionful correlation length must remain the same. However, the lattice spacing has increased by a factor of $b = 3$ (in this particular example). Thus the dimensionless correlation length transforms according to

$$\xi(x') = b^{-1}\xi(x), \tag{3.12}$$

where $x' = x^b$. This has the solution

$$\xi(x) = \frac{\text{const.}}{\ln x} = \frac{\text{const.}}{\ln \tanh K}, \tag{3.13}$$

which is the exact result for the one-dimensional Ising model. We see that ξ is finite, as expected, but that, as $T \to 0$, $\xi \propto e^{\text{const.}/T}$, so that it grows very large as the system approaches perfect ordering at $T = 0$.

Higher dimensions

The kind of analytic block spin renormalization group we have discussed above is no longer feasible for $d > 1$. Progress is possible by making various simplifying approximations, but, even then, the calculations rapidly become cumbersome and give little insight into the physics. Rather than pursue this subject in detail, then, we shall be content with making some general observations.

In one dimension it follows from (3.9) that at low temperature ($K \to \infty$) the renormalization group equation simplifies to $K' \sim K - \text{const.}$ This may be easily understood on physical grounds: at low temperature the spins in the blocks are almost always aligned in the same state. The interaction between adjacent blocks is mediated by their boundary spins (s_3 and s_4 in Figure 3.4). Thus we may write

$$K' \sim K \langle s_3 \rangle_{s_1'=1} \langle s_4 \rangle_{s_2'=1}, \tag{3.14}$$

where $\langle s_3 \rangle_{s_1'=1}$ is the magnetisation of the boundary spin, given that the block spin is $+1$. At low temperature, this is unity, so that $K' \sim K$.

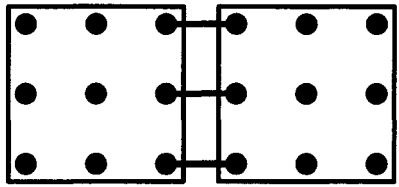


Figure 3.6. Neighbouring blocks in two dimensions.

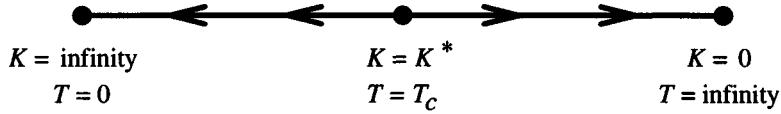


Figure 3.7. Schematic RG flow in the Ising model for $d > 1$.

Now consider the case of two neighbouring blocks in two dimensions, as illustrated in Figure 3.6. The interaction between neighbouring blocks is now mediated by three nearest neighbour bonds, so that, as $K \to \infty$, we expect that $K' \sim 3K$. In the case of d dimensions, with a length rescaling factor b, this generalises to

$$K' \sim b^{d-1}K \qquad \text{as} \qquad K \to \infty. \tag{3.15}$$

This has the important consequence that, for $d > 1$, $K' > K$, so that the zero-temperature fixed point at $K^{-1} = 0$ is locally stable. On the other hand, at high temperatures the system has to be in a paramagnetic phase, so the high-temperature fixed point must also be stable. Therefore, to the extent that we can think of the renormalization group flows as being unidimensional, there must exist an unstable fixed point $K = K^*$ in between them, as shown in Figure 3.7. This fixed point corresponds to the critical point. To see this, imagine calculating the correlation length using $\xi(K) = b\xi(K')$. Suppose that at some fixed high temperature, corresponding to a reduced coupling K_0, $\xi = \xi_0 = O(1)$. Starting at some $K < K^*$, it will take a certain number $n(K)$ of iterations of the renormalization group before we reach the vicinity of K_0. Then $\xi(K) = \xi_0 b^{n(K)}$. As K approaches K^*, the amount by which K changes under each iteration is initially very small, so that $n(K)$ becomes large. As $K \to K^*$, $n(K) \to \infty$, so that $\xi(K) \to \infty$,

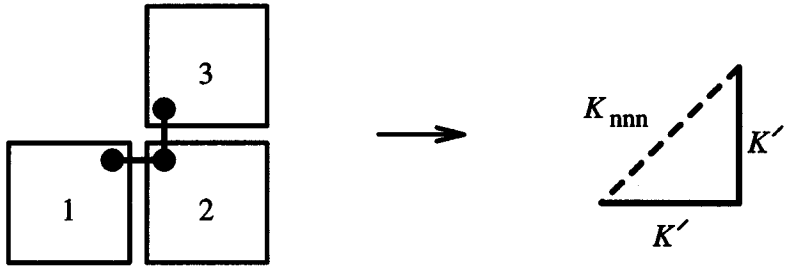


Figure 3.8. Generation of next-nearest neighbour coupling.

indicating a critical point.

In fact, knowing how K' depends on K close to the fixed point, we may calculate the critical exponent ν. Suppose that $K' = \mathcal{R}(K)$, where $K^* = \mathcal{R}(K^*)$. When $K - K^*$ is small, we may then write

$$K' \approx \mathcal{R}(K^*) + (K - K^*)\mathcal{R}'(K^*) = K^* + b^y(K - K^*), \quad (3.16)$$

which defines the quantity $y \equiv \ln \mathcal{R}'(K^*)/\ln b$. Now, close to the critical point, we expect from p.7 that $\xi(K) \sim A(K - K^*)^{-\nu}$. Using $\xi(K) = b\xi(K')$, it then follows that

$$A(K - K^*)^{-\nu} = bA(K' - K^*)^{-\nu} = bA\,[b^y(K - K^*)]^{-\nu}, \quad (3.17)$$

which is only possible if $\nu = 1/y$. This is an example of a general result (see Section 3.5) that critical exponents are given in terms of the *derivatives* of the renormalization group transformation at the fixed point. Indeed, we see from the above that we may trace the very existence of a critical exponent to the assumption of differentiability of the renormalization group transformation at the fixed point.

However, our analysis has, up to now, been lacking in one important respect, which does not arise in one dimension. Consider the effect of summing over the corner spin of block 2 in Figure 3.8: This couples to spins in blocks 1 and 3, so that it will generate an effective coupling between the respective block spins $1'$ and $3'$. On the blocked lattice, this will be a next-nearest neighbour coupling. In fact, as discussed earlier, all possible further neighbour couplings will now appear, and the one-dimensional picture of the renormalization group flows in Figure 3.7 is therefore a gross over-

simplification. However, it turns out that it is still the fixed points of the renormalization group transformation, now in the space of all possible couplings, which control the critical behaviour, and it is this feature which is responsible for the remarkable phenomenon of universality.

3.3 General theory

In this section we shall examine the consequences of the very general assumption that there exists a renormalization group fixed point in the space of all possible couplings. The transformation has the form $\{K'\} = \mathcal{R}(\{K\})$, where $\mathcal{R}$ will depend, in general, on the specific transformation chosen, and, in particular, on the length rescaling parameter b. Suppose there is a fixed point at $\{K\} = \{K^*\}$. As in the single variable case, we shall assume that $\mathcal{R}$ is differentiable at the fixed point, so that the renormalization group equations, linearised about the fixed point, are

$$K'_a - K^*_a \sim \sum_b T_{ab}(K_b - K^*_b), \tag{3.18}$$

where $T_{ab} = \partial K'_a/\partial K_b|_{K=K^*}$. Denote the eigenvalues of the matrix $\mathbf{T}$ by λ^i, and its left eigenvectors by $\{e^i\}$, so that

$$\sum_a e^i_a T_{ab} = \lambda^i e^i_b. \tag{3.19}$$

Note that we have no reason to suppose that $\mathbf{T}$ is symmetric, so that its left eigenvectors are not in general the same as the corresponding right eigenvectors. In fact, we are not even entitled to assume that the eigenvalues λ^i are real, but, as may be seen from the subsequent discussion, strange things would happen if they were not.†

Now define *scaling variables*‡ $u_i \equiv \sum_a e^i_a(K_a - K^*_a)$, which are linear combinations of the deviations $K_a - K^*_a$ from the fixed point

† In some random systems, the irrelevant eigenvalues, which correspond to corrections to scaling (see Section 3.6) may occur in complex conjugate pairs.

‡ See the discussion at the foot of p.53 on terminology.

which transform *multiplicatively* near the fixed point:

$$u_i' = \sum_a e_a^i (K_a' - K_a^*) = \sum_{a,b} e_a^i T_{ab} (K_b - K_b^*)$$
$$= \sum_b \lambda^i e_b^i (K_b - K_b^*) = \lambda^i u_i. \tag{3.20}$$

It is convenient to define the quantities y_i by $\lambda^i = b^{y_i}$. The y_i are called *renormalization group eigenvalues*, and will turn out to be related to the critical exponents. We may distinguish three cases:

- If $y_i > 0$, u_i is said to be *relevant*: repeated renormalization group iterations drive it away from its fixed point value.
- If $y_i < 0$, u_i is *irrelevant*: if we start sufficiently close to the fixed point, u_i will iterate towards zero.
- If $y_i = 0$, u_i is *marginal*. In this case, we cannot tell from the linearised equations whether u_i will move away from the fixed point or towards it. An example of this interesting case will be discussed later in Section 5.6.

Let us now consider a fixed point which has n relevant eigenvalues. For convenience, we may imagine the space near the fixed point as having n' dimensions in all (although strictly speaking this is infinite). There will then be $(n' - n)$ irrelevant eigenvalues, so, in the vicinity of the fixed point, there will be an $(n' - n)$-dimensional hypersurface of points attracted into the fixed point. Near the fixed point, this is just the linear space spanned by the irrelevant eigenvectors, but, by continuity, we expect this hypersurface to exist in some finite region around the fixed point. It is called the *critical surface*, since the long distance properties of each system corresponding to a point on this surface will be controlled by the fixed point, at which, by the same arguments as in Section 3.2, the correlation length will be infinite. Now the coupling constants K_a will depend in some complicated manner on the physical parameters like the temperature, pressure or magnetic field which the experimentalist may vary. We shall refer to these as 'knobs' which the experimentalist may adjust. In order to end up on the $(n'-n)$-dimensional surface attracted into this fixed point, she must therefore adjust exactly n knobs.

In the example of the ferromagnetic Ising model, two knobs must be adjusted (T and H) to bring the system to its critical

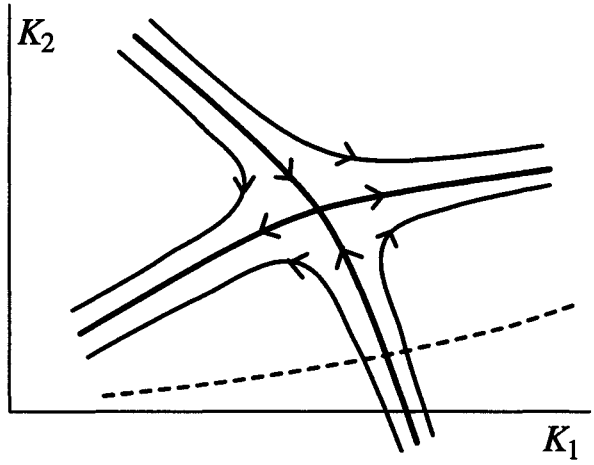


Figure 3.9. RG flows in a two-dimensional example.

point. The same is true of a simple fluid (e.g. pressure and temperature). Thus we expect the fixed point corresponding to this universality class to have just two relevant scaling variables. Since the critical point occurs at zero field in the Ising model, the corresponding fixed point must also occur in the subspace where all the odd couplings vanish. This means that the matrix $\mathbf{T}$ must be block diagonal, having no elements which connect the even and odd subspaces, and we may therefore classify all its eigenvectors as being either even or odd under the symmetry $s \to -s$. Thus one of the relevant variables must be temperature-like, and lie in the even subspace – this is called the *thermal* scaling variable – and the other, the *magnetic* scaling variable, must lie in the odd subspace.

In order to illustrate the significance of the irrelevant fields, let us restrict attention to flows in the even subspace, and, for simplicity, reduce the number of dimensions of this subspace to two, so that it is parametrised by (K_1, K_2). For example, K_1 might be the reduced nearest neighbour coupling, and K_2 the next-nearest neighbour coupling. There is then just one relevant and one irrelevant eigenvalue. The topology of the flows near the fixed point is shown in Figure 3.9. There is a one-dimensional curve of points attracted into the fixed point. This is the critical surface in this

toy example. The flows near the fixed point must have the hyperbolic form shown in the figure by continuity. Thus the critical surface, in this example, acts as a separatrix, dividing the region of points which flow to large values of the K_a (ultimately to zero temperature) from those flowing to small K_a (ultimately to the high-temperature fixed point at $K_a = 0$). In the model with nearest neighbour couplings only, as the temperature is varied we move along the axis $K_2 = 0$. The point where this line meets the critical surface then defines the critical reduced coupling K_{1c}, corresponding to the critical temperature of the nearest neighbour model, since points with $K < K_{1c}$ and $K > K_{1c}$ end up at the high- or low-temperature fixed points respectively. At $K = K_{1c}$, the renormalization group trajectories flow into the critical fixed point, which means that the long distance behaviour at the critical point is the same as that of the fixed point.

However, we may equally well consider a model with a next-nearest neighbour coupling in addition. Then, as we change the temperature, we move along some other curve in coupling constant space, indicated by the dashed line in Figure 3.9. The critical point of this model occurs where this line intersects the critical surface. But the large distance behaviour in this case will be similar to that in the case of a simple nearest neighbour coupling, because they are both controlled by the same fixed point.

This argument, suitably generalised to the case of an infinite-dimensional coupling constant space, is the simple explanation of the phenomenon of universality. A universality class consists of all those critical models which flow into a particular fixed point. To each universality class will correspond a different critical fixed point. However, in order to understand precisely which quantities are universal, we need to understand just what information the fixed point actually provides about the critical theory.

3.4 Scaling behaviour of the free energy

For definiteness, let us continue to consider the universality class of the critical short-range Ising model. As discussed in the previous section, there is a relevant thermal scaling variable u_t, with eigenvalue y_t, and a relevant magnetic scaling variable u_h, with eigenvalue y_h. In addition, there will be an infinite number of ir-

relevant variables $u_3, \ldots$ The critical point of the model in which we are interested will, in general, lie some finite distance away from the fixed point in coupling constant space. However, only a *finite* number of renormalization group iterations will be required to bring the renormalized theory to the vicinity of the fixed point, where the linearised version of the renormalization group equations is valid. The values u_i of the scaling fields at this point will therefore depend *analytically* on the deviations (t, h) of the original theory from its critical point. This is because the renormalization group transformation itself is analytic, and so therefore is also the result of a finite number of iterations. The relevant variables (u_t, u_h) must also vanish when $t = h = 0$, so that, by symmetry, they must have the form

$$u_t = t/t_0 + O(t^2, h^2) \tag{3.21}$$

$$u_h = h/h_0 + O(th), \tag{3.22}$$

where t_0 and h_0 are non-universal constants. Therefore, close to the critical point, we may take u_t and u_h to be proportional to t and h respectively.†

Recall that one of the properties of the renormalization group transformation is that it preserves the partition function:

$$Z = \mathrm{Tr}_s\, e^{-\mathcal{H}(s)} = \mathrm{Tr}_{s'}\, e^{-\mathcal{H}'(s')}. \tag{3.23}$$

Consider the *reduced free energy per site*, $f(\{K\}) \equiv -N^{-1} \ln Z$, as a function of the couplings $\{K\}$. Under renormalization, the couplings flow according to the renormalization group equations, but in addition a constant term $Ng(\{K\})$ is added to the free energy, as in (3.10). Thus

$$e^{-Nf(\{K\})} = e^{-Ng(\{K\})-N'f(\{K'\})}, \tag{3.24}$$

where $N' = b^{-d}N$ is the total number of blocks. This gives the fundamental transformation law for the free energy per site:

$$\boxed{f(\{K\}) = g(\{K\}) + b^{-d} f(\{K'\})} \tag{3.25}$$

Notice that the free energy, unlike the correlation length, transforms inhomogeneously under the renormalization group. However, if we are interested in extracting only the *singular* behaviour

† Note that for fixed points with more relevant variables, e.g. at a tricritical point (Section 4.1), there is no reason to suppose that the scaling fields are simply proportional to the experimentalist's 'knobs' unless there is some symmetry enforcing this.

of f, for the purpose, for example, of calculating the critical exponents, we may in fact drop the inhomogeneous g term. Physically this is because it originates from summing over the short wavelength degrees of freedom within each block, so that $g(\{K\})$ should be an analytic function of the K_a, even at the critical point. In this manner we obtain a *homogeneous* transformation law for the *singular part* of the free energy f_s

$$f_s(\{K\}) = b^{-d} f_s(\{K'\}). \tag{3.26}$$

Close to the fixed point, we may write this in terms of the scaling variables

$$f_s(u_t, u_h) = b^{-d} f_s(b^{y_t} u_t, b^{y_h} u_h) = b^{-nd} f_s(b^{ny_t} u_t, b^{ny_h} u_h), \tag{3.27}$$

where, for the time being, the irrelevant variables $u_3, \ldots$ are ignored. In the last expression, we have iterated the renormalization group n times. Since the variables u_t and u_h are growing under this iteration, we cannot make n too large, or the linear approximation to the renormalization group equations would eventually break down. So let us choose to halt the iteration at the point where $|b^{ny_t} u_t| = u_{t0}$, where u_{t0} is arbitrary but fixed, and sufficiently small so that the linear approximation is still valid. Solving this equation for n, we then find, after a little algebra, that

$$f_s(u_t, u_h) = |u_t/u_{t0}|^{d/y_t} f_s\left(\pm u_{t0}, u_h |u_t/u_{t0}|^{-y_h/y_t}\right). \tag{3.28}$$

Rewriting this in terms of the reduced physical variables t and h, we see that u_{t0} may be incorporated into a redefinition of the scale factor t_0, and that

$$f_s(t, h) = |t/t_0|^{d/y_t} \Phi\left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}}\right), \tag{3.29}$$

where Φ is a *scaling function.*† This function might appear to depend on u_{t0}, but since the left hand side of (3.29) cannot, this is illusory, and, in fact, such scaling functions turn out to be *universal.* The only dependence on the particular system is through the *scale factors* t_0 and h_0.

3.5 Critical exponents

From the scaling law (3.29) for the singular part of the free energy follow all the thermodynamic exponents:

† Note that there are, in fact, different scaling functions for $t > 0$ and $t < 0$.

- Specific heat $\partial^2 f/\partial t^2|_{h=0} \propto |t|^{d/y_t - 2}$, so that

$$\boxed{\alpha = 2 - d/y_t} \tag{3.30}$$

- Spontaneous magnetisation $\partial f/\partial h|_{h=0} \propto (-t)^{(d-y_h)/y_t}$, so that

$$\boxed{\beta = \frac{d - y_h}{y_t}} \tag{3.31}$$

- Susceptibility $\partial^2 f/\partial h^2|_{h=0} \propto |t|^{(d-2y_h)/y_t}$, so that

$$\boxed{\gamma = \frac{2y_h - d}{y_t}} \tag{3.32}$$

- To get δ we must work a little harder: we have

$$M = \frac{\partial f}{\partial h} = |t/t_0|^{(d-y_h)/y_t}\, \Phi'\left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}}\right). \tag{3.33}$$

When inverted to express h as a function of M, this is called the scaling form of the equation of state (Widom scaling). For M to have a finite limit as $t \to 0$, $\Phi'(x)$ must behave like $x^{d/y_h - 1}$ as $x \to \infty$. Thus, at $t = 0$, $M \propto h^{d/y_h - 1}$, or

$$\boxed{\delta = \frac{y_h}{d - y_h}} \tag{3.34}$$

We see that the four principal thermodynamic exponents are given in terms of the two renormalization group eigenvalues y_t and y_h. This means that there must exist *scaling relations* between them. Examples are

$$\alpha + 2\beta + \gamma = 2, \tag{3.35}$$

$$\alpha + \beta(1 + \delta) = 2. \tag{3.36}$$

These relations, among others, were postulated before the advent of the renormalization group. Many of them may be proved rigorously as inequalities. Although the relations between the critical exponents and the renormalization group eigenvalues were established above for the case of the Ising universality class, similar equations hold for any universality class with a single relevant thermal eigenvalue and a relevant symmetry-breaking field. However, for more complicated cases, describing for example multicritical points (see Sections 4.1 and 4.2), the corresponding relations are more complex and should be derived from first principles as above.

Role of the rescaling factor b

In the previous section, we saw how the various thermodynamic exponents are related to the eigenvalues b^{y_i} of the matrix T_{ab} of derivatives of the renormalization group transformation at the fixed point. Since the length rescaling factor b enters explicitly into this calculation, one might legitimately ask whether it plays a role in the final values for the exponents. The answer is, of course, that it cannot, since the exponents are properties of the system under consideration, rather than the particular renormalization group transformation which is applied. In fact, the renormalization group transformation contains b implicitly, and it must therefore happen that, *if* the transformation is carried out exactly, the physical observables such as the exponents and scaling functions are independent of b. In the few cases where an exact renormalization group solution is available (for example in one dimension) or when it may be cast in the form of a controlled approximation (as in the ϵ-expansion), the independence of the final results of the details of the transformation may indeed be verified. In the case of block spin transformations, however, which usually demand uncontrolled approximations, the *approximate* values obtained for the exponents do exhibit weak dependence on b and other details of the transformation.

Although for a block spin transformation the possible values of the rescaling factor b are strongly limited by the requirement that the blocked lattice should have the same structure as the original, as we shall see later there are other forms of the renormalization group in which b may be arbitrary. In these cases it is often helpful to consider the limit of an *infinitesimal* transformation, when $b = 1 + \delta\ell$, with $\delta\ell \ll 1$. In this case, the couplings will also transform infinitesimally

$$K_a \to K_a + (dK_a/d\ell)\delta\ell + O(\delta\ell^2), \tag{3.37}$$

and the renormalization group equations take the differential form

$$dK_a/d\ell = -\beta_a(\{K\}), \tag{3.38}$$

where the functions β_a are called the renormalization group beta functions. (The minus sign in (3.38) is included so as to make contact with the conventional definition of the beta function in quantum field theory, $\beta_a = \kappa\partial K_a/\partial\kappa$, where κ is a wave number,

rather than a length scale.) In this infinitesimal form of the renormalization group, the fixed points now correspond to the *zeroes* of the beta functions. The matrix of derivatives at the fixed point is now $T_{ab} = \delta_{ab} + (\partial\beta_a/\partial K_b)\delta\ell$, with eigenvalues $(1+\delta\ell)^{y_i} \sim 1+y_i\delta\ell$. Hence the y_i are simply the eigenvalues of the matrix $-\partial\beta_a/\partial K_b$, evaluated at the zero of the beta functions. These infinitesimal renormalization group transformations will play a central role in Chapter 5.

3.6 Irrelevant eigenvalues

Suppose that we have an irrelevant scaling variable u_3, with eigenvalue $y_3 < 0$. If it is in the even subspace, we may assume that its initial value depends analytically on t and h, and is therefore of the form

$$u_3 = u_3^0 + at + bh^2 + \cdots, \tag{3.39}$$

where $a, b, \ldots$ are constants. Unlike the case of a relevant variable, however, we may not assume that $u_3^0 = 0$. Close to the critical point, we may initially ignore the higher order terms and set $u_3 = u_3^0$, which has some non-universal value. Repeating our calculation in Section 3.4 of the free energy, we now find that

$$f_s(t,h) \sim |t|^{d/y_t}\, \Phi\left(h|t|^{-y_h/y_t}, u_3^0|t|^{|y_3|/y_t}\right). \tag{3.40}$$

Since $u_3^0|t|^{|y_3|/y_t}$ is small as $t \to 0$, and the right hand side represents the free energy evaluated away from the critical point, it seems reasonable to assume that it is an analytic function of its arguments, and that, in particular, we expand it in a Taylor series in $u_3^0|t|^{|y_3|/y_t}$. Taking $h = 0$ for clarity, we then find that

$$f_s = |t|^{d/y_t}\left(A_1 + A_2 u_3^0|t|^{|y_3|/y_t} + \cdots\right), \tag{3.41}$$

where A_1, A_2, *etc.* are non-universal constants. We see that the leading effect of such irrelevant variables is to give rise to *correction to scaling* terms.† Since in a real system the coefficient of these terms may be quite large, it may be difficult to observe the true asymptotic exponents except very close to the critical point. In that case, a fit to the data which does not include such correction

† These are sometimes called ***confluent singularities***, usually in the context of extracting critical behaviour from high-temperature expansions.

to scaling terms may erroneously lead to the conclusion that the exponents extracted in this way are non-universal. On the other hand, fitting data with the correction to scaling terms included introduces many more parameters, and thus requires very high quality data to obtain a meaningful fit.

As well as the non-analytic correction to scaling terms of the type shown in (3.41), there are also corrections that come from the higher order dependence of the starting values of the u_i on t and h. Typically, they lead to corrections which are down by relative integral powers of t and h on the leading terms, and are thus called analytic corrections. Since the exponent $|y_3|/y_t$ is around 0.5 for many three-dimensional systems, the non-analytic corrections are more important close to the critical point.

Implicit in the above discussion was the assumption that the Taylor expansion of the right hand side of (3.41) exists, and the limit $u_3 \to 0$ is smooth and well-defined. However, there are situations, several of which will arise later (see Sections (4.4, 8.4, 9.4)), when this is not true. In such a case, u_3 is referred to as a *dangerous* irrelevant variable, and we need more information about the dependence on u_3 to infer from (3.41) the true behaviour of the free energy in the critical region.

3.7 Scaling for the correlation functions

The transformation law for the free energy (3.25) relied only on the property of the renormalization group that it preserves the partition function. However, as we have seen in Section 3.1, the renormalization group in fact preserves the whole probability measure of the long wavelength degrees of freedom, and therefore analogous transformation laws should apply to the large distance behaviour of the correlation functions. This is indeed the case. As an example we shall consider the spin-spin 2-point correlation function in the Ising model, defined by

$$G(r_1 - r_2, \mathcal{H}) \equiv \langle s(r_1)s(r_2)\rangle_{\mathcal{H}} - \langle s(r_1)\rangle_{\mathcal{H}}\langle s(r_2)\rangle_{\mathcal{H}}, \qquad (3.42)$$

where we have emphasised the dependence of G on the parameters in the hamiltonian $\mathcal{H}$. We may also obtain G by adding a non-uniform magnetic field to the hamiltonian $\mathcal{H} \to \mathcal{H} - \sum_r h(r)s(r)$

and differentiating the free energy with respect to $h(r)$:

$$G(r_1 - r_2) = \frac{\partial^2}{\partial h(r_1)\partial h(r_2)} \ln Z\{h\}\Big|_{h(r)=0}. \tag{3.43}$$

We now suppose that $h(r)$ varies significantly only over distances much large than the size ba of the blocks, and imagine applying the same type of block spin renormalization group as in Section 3.1. If the hamiltonian $\mathcal{H}$ (which is close to the fixed point hamiltonian $\mathcal{H}^*$, since we are interested in the critical region) contains only short-range interactions, then, in performing the block spin transformation in one region, around r_1 say, we may effectively ignore the fact that $h(r)$ is actually slowly varying, and assume that it transforms in the same manner as would a weak *uniform* field $h = h(r_1)$. According to this argument, the renormalized hamiltonian is therefore of the same form

$$\mathcal{H}'(s') - \sum_{r'} h'(r')s'(r'), \tag{3.44}$$

where $h'(r') = b^{y_h}h(r)$. Since the renormalization group preserves the entire partition function, however,

$$\frac{\partial^2 \ln Z'(h')}{\partial h'(r_1')\partial h'(r_2')} = \frac{\partial^2 \ln Z(h)}{\partial h'(r_1')\partial h'(r_2')}. \tag{3.45}$$

Let us examine the meaning of each side of this equation. The left hand side is just the correlation function of the block spins in the ensemble defined by the renormalized hamiltonian $\mathcal{H}'$. However, in units of the lattice spacing, the distance between the points has been reduced by a factor of b. Thus the left hand side is simply $G((r_1 - r_2)/b, \mathcal{H}')$. The right hand side is more tricky. Making an infinitesimal local change $h'(r_1') \to h'(r_1') + \delta h'(r_1')$ within block number 1 corresponds to changing *all* the fields $h(r_i)$ acting on the spins within this block, by an amount $\delta h(r_i) = b^{-y_h}\delta h'(r_1')$. Thus the right hand side is

$$b^{-2y_h}\langle(s_1^{(1)} + s_2^{(1)} + \cdots)(s_1^{(2)} + s_2^{(2)} + \cdots)\rangle_{\mathcal{H}}, \tag{3.46}$$

where the spins in blocks 1 and 2 are labelled by $s_i^{(1)}$ and $s_i^{(2)}$ respectively, and the subscript denotes that this correlation function is to be evaluated with respect to the original hamiltonian $\mathcal{H}$. Since there are b^d spins in each block, (3.46) may be expanded as a sum of b^{2d} two-point correlations. If $|r_1 - r_2|$ is much larger

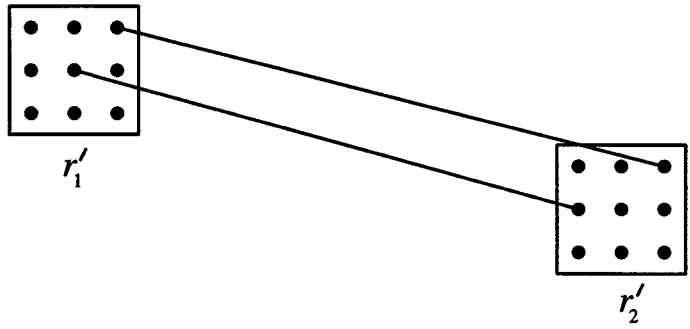


Figure 3.10. Correlations between spins in two distant blocks.

than b, all these correlation functions are, however, numerically almost the same (see Figure 3.10). The transformation law for the correlation function, close to the fixed point, is therefore

$$G((r_1 - r_2)/b, \mathcal{H}') = b^{2(d-y_h)} G(r_1 - r_2, \mathcal{H}). \tag{3.47}$$

If the interactions are isotropic (respecting, for example, the rotational symmetries of the lattice), then, at large enough distances, the 2-point correlation function in fact depends only on the distance $r = |r_1 - r_2|$ and not on the relative orientation of the two points. This is because scaling fields which break the full rotational symmetry down to that of the lattice are irrelevant.

Setting the uniform magnetic field $h = 0$ for clarity, near the critical point

$$G(r, t) = b^{-2(d-y_h)} G(r/b, b^{y_t} t). \tag{3.48}$$

We may now iterate this equation n times, as we did for the free energy, stopping at a point where $b^{ny_t}(t/t_0) = 1$. After a little algebra, it follows that the correlation function has the scaling form

$$G(r, t) = |t/t_0|^{2(d-y_h)/y_t} \Psi \left(r/|t/t_0|^{-1/y_t} \right). \tag{3.49}$$

For sufficiently large r, we expect G to decay as $e^{-r/\xi}$, since this also corresponds to large t and the region where the mean field result of Section 2.3 should apply. From (3.49) we then identify the correlation length $\xi \propto |t|^{-1/y_t}$, so that

$$\boxed{\nu = 1/y_t} \tag{3.50}$$

At the critical point $t = 0$ we should instead iterate the transformation law (3.48) until $r/b^n = O(r_0)$, where r_0 is some fixed distance, much larger than a or the range of the interaction, so that all the approximations made above are still valid. We then see that $G(r) \propto r^{-2(d-y_h)}$, so that

$$\boxed{\eta = d + 2 - 2y_h} \tag{3.51}$$

The exponents ν and η, related to the spin-spin correlation function, are therefore also given in terms of the relevant renormalization group eigenvalues y_t and y_h. They are thereby also related to the thermodynamic exponents, by scaling relations such as

$$\alpha = 2 - d\nu \tag{3.52}$$

$$\gamma = \nu(2 - \eta). \tag{3.53}$$

Results of this kind clearly require further assumptions than went into the scaling relations such as (3.35, 3.36) for the thermodynamic exponents. A crucial input is that the fixed point hamiltonian be sufficiently short range. For long range interactions (see Section 4.3), for example, (3.53) fails.

(3.52) is an example of a *hyperscaling* relation, since it connects the singularity in the specific heat with the behaviour of a correlation length, which may be inferred, for example, from the spin correlation function. This relation may fail when a dangerous irrelevant variable (see the previous section) influences the scaling form of the free energy but not of the correlation functions. This is what happens above the upper critical dimension (see Section 5.4), and in some problems, even below it. Examples are the random field Ising model (Section 8.4), and the branched polymer problem (Section 9.4).

3.8 Scaling operators and scaling dimensions

The above discussion of the spin-spin correlation function of the Ising model may readily be generalised to an arbitrary correlation function. Near a general fixed point, the scaling variables u_i are linear combinations of the deviations $K_a - K_a^*$ of the original couplings from their fixed point values. Each of these couples to a unique possible interaction term S_a in the hamiltonian. Each S_a may be expressed in terms of the fundamental degrees of freedom

of the problem (for the Ising model, the spins $s(r)$). For example, in the Ising model, the set $\{S_a\}$ would be expressible as linear combinations of arbitrary products of spins on different sites. However, the assumption that only short-range couplings are important means that these objects should be *local*, in some sense to be made more precise. It has become common to call these composite objects *operators*. The reason is that, when a suitable continuum limit is taken, the statistical mechanics model is formally identical to a quantum field theory, in which these quantities become operators which may represent observables in the sense of quantum mechanics. However, it is important to realise that in this book, except when so stated explicitly, these 'operators' are commuting quantities.†

Given a complete set of operators S_a we may form suitable linear combinations ϕ_i, called *scaling operators*, coupling uniquely to each of the scaling fields u_i, so that

$$\sum_i u_i \phi_i = \sum_a (K_a - K_a^*) S_a. \tag{3.54}$$

It is then straightforward to generalise the argument of the previous section to show that, as $|r_1 - r_2| \to \infty$, $\langle \phi_i(r_1)\phi_i(r_2)\rangle \propto |r_1 - r_2|^{-2x_i}$, where

$$\boxed{x_i = d - y_i} \tag{3.55}$$

This equation, which relates the renormalization group eigenvalue of a scaling variable to the behaviour at the fixed point of the two-point correlation function of the operator to which it couples, is one of the most fundamental and general results of the renormalization group. The quantity x_i is called the *scaling dimension* of the scaling operator ϕ_i. The relation (3.55) may be understood if we assume that it is possible to take the continuum limit of the hamiltonian, in such a way that

$$\sum_i u_i \sum_r \phi_i(r) \to \sum_i u_i \int \phi_i(r) \frac{d^d r}{a^d}, \tag{3.56}$$

where a^d is the volume of the unit cell. If, under a renormalization group transformation where $a \to ba$ and $u_i \to b^{y_i} u_i$, we demand

† An alternative terminology is to call the u_i scaling 'fields', and the ϕ_i scaling 'densities'. This is, however, especially confusing when used in the context of quantum field theory, and we shall avoid it.

that the partition function be invariant, this may then be ensured by requiring that $\phi_i(r) \to b^{x_i}\phi_i(r)$, with x_i given by (3.55).

As an example, consider the local energy density $E(r)$, which, for the Ising model, is the product $s(r_1)s(r_1')$ of neighbouring spins. It has a scaling dimension $x_E = d - y_t = d - \nu^{-1}$. Therefore, at the critical point, its two-point correlation function decays as

$$\langle E(r_1)E(r_2)\rangle \sim |r_1 - r_2|^{-2d+2/\nu}, \tag{3.57}$$

which is confirmed by exact results in two dimensions. Actually, in writing results such as (3.57) we should remember that operators like $E(r)$ are not themselves scaling operators, but only linear combinations thereof. Only *scaling* operators have a pure power behaviour for their correlation functions. The lattice energy operator will, in general, be a combination of all possible scaling operators which transform in the same way under the symmetries of the fixed point hamiltonian. The most relevant of these, and hence the one with the smallest scaling dimension, will be the operator ϕ_t, whose 2-point function, at the fixed point, will have the pure power-law form (3.57). In general, for finite separations, (3.57) should be replaced by †

$$\langle E(r_1)E(r_2)\rangle = \sum_{i,j} \frac{A_{ij}}{|r_1 - r_2|^{x_i+x_j}}. \tag{3.58}$$

Even this is not the full story, since, in general, a system at its critical point is not at the fixed point, but, as discussed in Section 3.6, its hamiltonian differs by irrelevant operators. Just as for the free energy, these lead to correction to scaling terms which now show up as corrections to (3.58) of the form $|r_1 - r_2|^{-x_i-x_j-\sum_k' |y_k|}$, where the sum is over the eigenvalues of some subset of irrelevant scaling fields.

The usefulness of the concept of scaling dimension is not restricted to the two-point correlations. For example if we consider an N-point correlation function, the same kind of arguments that led to (3.55) now imply that this has the homogeneity property

$$\langle \phi_1(r_1)\phi_2(r_2)\ldots\phi_N(r_N)\rangle = R^{-x_1-\cdots-x_N}\langle \phi_1(r_1/R)\phi_2(r_2/R)\ldots\phi_N(r_N/R)\rangle, \tag{3.59}$$

† In fact, conformal invariance (see Section 11.2) implies that the terms with $i \neq j$ vanish in this sum.

although, unlike the case $N = 2$, this (together with translational and rotational invariance) is not sufficient to fix its actual form.

3.9 Critical amplitudes

In the previous sections we saw how the various critical exponents are directly related to the renormalization group eigenvalues at the corresponding fixed point. Since this fixed point controls the critical behaviour of all systems in a given universality class, it follows that the critical exponents are universal. However, many other quantities are universal besides the exponents. In this section we focus on the amplitudes which multiply the power law singularities in thermodynamic quantities near the critical point. It is necessary to go back to the inhomogeneous transformation law for the free energy (3.25). Iterating this n times

$$f(\{K\}) = \sum_{j=0}^{n-1} b^{-jd} g(\{K^{(j)}\}) + b^{-nd} f(\{K^{(n)}\}), \tag{3.60}$$

where $\{K^{(j)}\}$ is the jth iterate of $\{K\}$. Now take the limit $n \to \infty$. Since all initial values of the $\{K\}$ in the same phase ultimately iterate into the same stable fixed point $\{K^{(\infty)}\}$, the second term on the right hand side tends to zero. The first term is a weighted sum of the values of $g(\{K\})$ at points along the renormalization group trajectory which leads from the starting value of $\{K\}$ up to this stable fixed point (see Figure 3.9). For starting points sufficiently close to the critical surface, this trajectory (almost) breaks into two pieces: one which closely approximates a trajectory along the critical surface and which ends up near the critical fixed point; and one starting from near the vicinity of the critical fixed point and ending at the stable fixed point. This latter part closely approximates the *unique* trajectory leading to the stable fixed point which, if the arrows were reversed, would arrive at the critical fixed point. We may call this the outflow trajectory.† The sum in (3.60) may thus be broken into two pieces accordingly. The first contribution may be shown to give only correction to scaling terms

† It is the subspace of all models which correspond to renormalizable continuum quantum field theories, where all the irrelevant couplings have been set to zero.

in the free energy (as expected from the discussion in Section 3.6). To evaluate the second term, we may proceed as follows.

Define a coordinate $\tilde{u}_t$ along the outflow trajectory. Close to the critical fixed point, we may take $\tilde{u}_t \sim u_t$, the thermal scaling variable, but, further away, they will deviate from each other due to the curvature of the trajectory. However, in general, we may define $\tilde{u}_t$ so that it is a *nonlinear scaling variable* (see Ex. 3.5), which means that it transforms homogeneously under the renormalization group, $\tilde{u}_t \to b^{y_t}\tilde{u}_t$, for *all* values of $\tilde{u}_t$, not just close to the critical fixed point. Thus, the contribution to the free energy from the outgoing part of the trajectory, which will turn out to give rise to the leading singular behaviour as $t \to 0$, may be written

$$f(t) \sim \sum_{j=0}^{\infty} b^{-jd} g(b^{jy_t}\tilde{u}_t), \tag{3.61}$$

where $\tilde{u}_t = t/t_0$. Since we are interested in extracting the behaviour as $t \to 0$, it is permissible to replace the sum over j by an integration, and to change to the integration variable $s \equiv b^{jy_t}(t/t_0)$, whence

$$f(t) \sim \frac{\tilde{u}_t^{d/y_t}}{y_t \ln b} \int_{u_t}^{\infty} s^{-d/y_t - 1} g(s) ds. \tag{3.62}$$

All we need assume about the function $g(s)$ is that it is analytic, as was argued in Section 3.4, and that it and all its derivatives approach zero sufficiently fast as $s \to \infty$.

Superficially the right hand side of (3.62) appears to exhibit the $\tilde{u}_t^{d/y_t} = t^{2-\alpha}$ behaviour expected of the singular part of the free energy, as argued in Section 3.5. However, this conclusion is valid only if it is permissible to set the lower limit of the integration to zero, which is, in general, not the case since $g(0) \neq 0$ and the integral would diverge. It is necessary first to integrate by parts, in order to increase the power of s in the integrand. After p such integrations, the contribution from the lower limit is of the form $\tilde{u}_t^{d/y_t}\, \tilde{u}_t^{-d/y_t + p - 1} g^{(p-1)}(\tilde{u}_t)$, which is analytic in t. If we integrate by parts a sufficient number of times so that $p > d/y_t$, the remaining

integral is of the form†

$$\int_{u_t}^{\infty} s^{-d/y_t+p-1}\, g^{(p)}(s)\, ds, \tag{3.63}$$

in which is it now permissible to set the lower limit to zero (the corrections to this may once again be shown to be analytic in t).

The conclusion is that the contribution to the free energy near $t = 0$ from the outgoing part of the trajectory is of the form

$$f(t) \sim A_{>,<} |t|^{2-\alpha} + \text{terms analytic in } t, \tag{3.64}$$

where the subscripts $(>, <)$ on the amplitude A indicate that we expect it to be different for $t > 0$ and $t < 0$. This is because, in each case, the renormalization group trajectory along which g is integrated away from the critical fixed point is different. The amplitudes $A_{>,<}$ consist, apart from trivial factors, of a *universal* integral $\int_0^\infty s^{-d/y_t+p-1} g^{(p)}(s) ds$, and the non-universal scale factor $t_0^{-2+\alpha}$. This latter factor arises from the renormalization of u_t which occurs in the first part of the renormalization group flow from the critical theory into the vicinity of the critical fixed point. It is therefore the same no matter what the sign of t. We conclude that, although the individual amplitudes $A_{>,<}$ are not universal, their *ratio* $A_>/A_<$ is.

This is just one example of a universal amplitude ratio. In general, combinations of critical amplitudes in which the non-universal scale factors like $t_0, h_0, \ldots$ do not enter are expected to be universal. These non-universal factors determine how the non-linear scaling variables, which describe the position along the outflow trajectory, are related to the physical 'knobs' which the experimentalist may adjust. Once on the outflow trajectory, however, everything is universal.

Another important example of a universal amplitude combination is the quantity $f_s \xi^d$, where ξ is the correlation length. Since f_s is the singular part of the free energy per unit volume, this combination is the free energy per correlation volume. In terms of the exponents defined earlier, it should scale as $t^{2-\alpha} \cdot t^{-d\nu}$, and so should be independent of t if the hyperscaling relation (3.52) holds. The universality of the numerical value of $f_s \xi^d$ is therefore a stronger statement of hyperscaling.

† An interesting situation arises when d/y_t is an integer (see Ex. 3.6).

The above arguments show, in general, that such universal combinations are in principle calculable from the renormalization group, although they depend on properties of the renormalization group flows not just at the critical fixed point, but on more global quantities.

Anisotropic scaling

In deriving (3.48) we assumed that the two-point correlation function depends, at least at large distances, only on the magnitude $r_{12} = |r_1 - r_2|$ of the separation between the two points. As discussed there, this is certainly a correct assumption if the underlying lattice model is sufficiently isotropic. If this is not the case, however, various other possibilities may occur. These may be understood by examining the Fourier transform of the exchange interaction $J(r-r')$, which, for small wave numbers $\mathbf{k} = (k_x, k_y, \ldots)$, has the form

$$\tilde{J}(\mathbf{k}) = J(1 - \sum R_i^2 k_i^2 + O(k^4)). \tag{3.65}$$

If none of the R_i^2 vanishes, we may simply perform a rescaling of the coordinates so that, at least to $O(k^2)$, (3.65) is isotropic. Although this argument is strictly valid only within mean field theory, since, as will be argued in Chapter 4, the higher powers of k are irrelevant at the isotropic fixed point, it in fact continues to hold when the fluctuations are included.†

However, a much more severe kind of anisotropy may arise if one of the R_i^2 in (3.65) vanishes. For example, $\tilde{J}$ may have the expansion for low wave number

$$\tilde{J}(k) = J(1 - R^2 k_\perp^2 - \lambda R^4 k_{||}^4 + \cdots), \tag{3.66}$$

where $\mathbf{k} = (k_{||}, \mathbf{k}_\perp)$, and λ is dimensionless constant. Such a behaviour arises, for example, at the *Lifshitz point* in a system with competing ferromagnetic nearest neighbour and antiferromagnetic next-nearest neighbour interactions in one particular direction. In this case, it is clearly not possible to remove the anisotropy by a simple rescaling of the coordinates. In mean field

† However, the quantitative rescaling required to render the model isotropic does depend on the value of these irrelevant terms, and so is not simply given by the mean field result.

theory, the Fourier transform of the correlation function for such a system has the form

$$\tilde{G}(k) \propto (\lambda k_{\|}^4 + k_{\perp}^2 + \xi^{-2})^{-1}. \tag{3.67}$$

However, when the fluctuations are included, there is a further complication in that λ is, in general, renormalized, since there is no symmetry which protects this. Because of the intrinsic anisotropy, it no longer makes sense to rescale all distances, both in the $\perp$ and $\|$ subspaces, by the same constant b. Instead, one should rescale only $r_{\perp}$, for example, and allow λ to be renormalized in such a way that the low wave-number physics is preserved. A little thought then shows that, on dimensional grounds, the renormalization group equation for λ must take the form

$$\lambda' = \lambda f(\{K\}), \tag{3.68}$$

where f depends on all the other dimensionless couplings. At the fixed point, then, we expect $\lambda' = f(\{K^*\})\lambda \equiv b^{-2/z}\lambda$, defining the *anisotropic scaling exponent* z. This shows up in several ways. For example, the scaling form (3.49) is replaced by

$$G(r_{\|}, r_{\perp}, t) = \xi_{\perp}^{-(d-2+\eta)} \Psi\left(r_{\perp}/\xi_{\perp}, r_{\|}/r_{\perp}^z\right), \tag{3.69}$$

where $\xi_{\perp} \propto |t|^{-\nu_{\perp}}$ is the correlation length in the $\perp$ directions. This scaling form then implies that the correlation length in the other direction diverges like $|t|^{-\nu_{\|}}$, where $\nu_{\|} = z\nu_{\perp}$. The correlation lengths in the two directions therefore exhibit quite different scaling behaviour. Note that, in the mean field approximation, we have $\nu_{\perp} = \frac{1}{2}$, $\nu_{\|} = \frac{1}{4}$ and $z = \frac{1}{2}$ in this example. (3.69) is the general form for a two-point correlation function exhibiting *anisotropic scaling*. Apart from the example of a Lifshitz point, we shall see it in the problem of directed percolation (p.200), and, when the $\|$ direction is interpreted as time, in quantum critical behaviour (p.76) and critical dynamics (p.192).

Exercises

3.1 The one-dimensional Ising model in a magnetic field has the reduced hamiltonian $\mathcal{H} = -K\sum_j s_j s_{j+1} - h\sum_j s_j$. By summing over every other spin, show that you can define a renormalization group transformation with $b = 2$. It is useful to express this in terms of the variables $x = e^{-2K}$ and $y = e^h$.

Sketch the flows in the (x, y) plane and indicate the fixed points.

3.2 The one-dimensional three-state Potts model is defined as follows: at each site j is a 'spin' which may take the values 1, 2 or 3. The reduced hamiltonian is $\mathcal{H} = -K \sum_j \delta_{s_j, s_{j+1}}$. Using the same decimation procedure with $b = 2$ as above, find the renormalization group equation and show that there are no non-trivial fixed points, as expected in one dimension.

3.3 By solving the one-dimensional Ising model exactly (e.g. by the transfer matrix method – see the references by Stanley and by Baxter in the Bibliography) show that the combination $f_s \xi$ is constant in the low temperature limit when $\xi \to \infty$, and calculate its numerical value. Do the same calculation for a spin 1 Ising model, with the same form for the hamiltonian but with spins s_j taking the values $0, \pm 1$, and verify that the amplitude combination is indeed universal.

3.4 There is an amplitude corresponding to each of the critical exponents defined for a ferromagnet on p.7. List as many combinations of these as you can which should be universal according to the arguments of Section 3.7.

3.5 Suppose that the infinitesimal renormalization group equations have the form $dg_i/d\ell = -\beta_i(\{g\})$ where the right hand sides have a perturbative expansion of the form $\beta_i = -y_i g_i + \sum_{jk} c_{ijk} g_j g_k + \cdots$. Show that, in general, it is possible to define, order by order in the gs, *non-linear* scaling variables $g_i' \equiv g_i + \sum_{jk} d_{ijk} g_j g_k + \cdots$ so that the renormalization group equations simplify to $dg_i'/d\ell = y_i g_i'$ exactly, with no higher order terms. Under what circumstances does such a transformation fail (to the order stated)? If there was a non-trivial fixed point at some finite value of the g_i in terms of the old variables, what has happened to it in terms of the non-linear scaling variables?

3.6 Show that when d/y_t is an integer, the free energy has a singularity of the form $|t|^{d/y_t} \ln |t|$. What can you say about the ratio of the amplitudes of this singular behaviour for $t > 0$ and $t < 0$?

3.7 Show that the contribution of the inflow part of the trajectory in the calculation of the free energy using (3.60) leads to the expected correction to scaling terms.