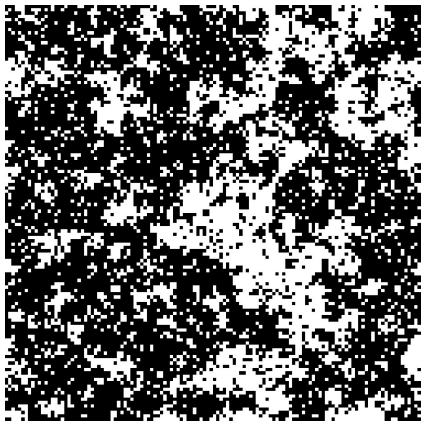
PH40045: Advanced Topics

Critical Point Phenomena

Lecturer: Nigel Wilding

November 2003



Literature

One motivation for supplying you with lecture notes for this course is the absence of a wholly ideal text book. However, it should be stressed that the notes are not fully comprehensive and should be regarded as the 'bare bones' of the course, to be fleshed out via your own reading and supplementary note taking. To this end perhaps the most appropriate textbook is

- *Statistical Mechanics of Phase Transitions*, by J.M. Yeomans (Oxford). There are 5 copies in the library, shelfmark 536.71 Yeo.

You might also wish to dip into the introductory chapters of the following more advanced texts

- *The Critical Point* C. Domb, (Taylor and Francis). Shelfmark 536.71
- *Lectures on Phase Transitions and the Renormalization Group*, N. Goldenfeld (Addison Wesley) Shelfmark 536.71 Gol.
- *The Theory of Critical Phenomena,* J.J. Binney, N.J. Dowrick, A.J. Fisher and M.E.J. Newman; (Oxford).

1 Introduction

A wide variety of physical systems undergo rearrangements of their internal constituents in response to the thermodynamic conditions to which they are subject. Two classic examples of systems displaying such phase transitions are the ferromagnet and fluid systems. As the temperature of a ferromagnet is increased, its magnetic moment is observed to decrease smoothly, until at a certain temperature known as the critical temperature, it vanishes altogether.

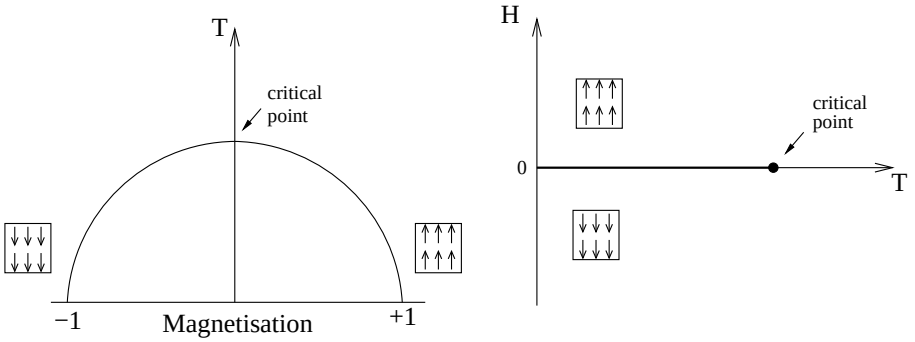


Figure 1: Phase diagram of a simple magnet (schematic)

Similarly, a change of state from liquid to gas can be induced in a fluid system (though not in an ideal gas) simply by raising the temperature. Typically the liquid-vapour transition is abrupt, reflecting the large density difference between the states either side of the transition. However the abruptness of this transition can be reduced by applying pressure. At one particular pressure and temperature the discontinuity in the density difference between the two states vanishes. These conditions of pressure and temperature serve to locate the critical point for the fluid.

In the vicinity of a critical point a system will exhibit a variety of remarkable effects known collectively as critical phenomena. Principal among these effects is the divergence of thermal response functions such as the specific heat and the fluid compressibility or magnetic susceptibility. It transpires that the origin of the singularities in these quantities can be traced to

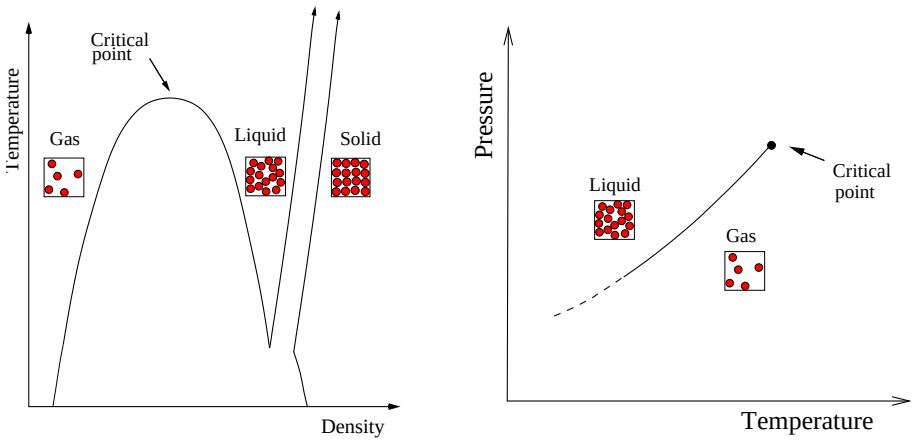


Figure 2: Phase diagram of a simple fluid (schematic)

large-length-scale co-operative effects between the microscopic constituents of the system. The recalcitrant problem posed by the critical region is how best to incorporate such collective effects within the framework of a rigorous mathematical theory that affords both physical insight and quantitative explanation of the observed phenomena. This matter has been (and still is!) the subject of intense theoretical activity.

The importance of the critical point stems largely from the fact that many of the phenomena observed in its vicinity are believed to be common to a whole range of apparently quite disparate physical systems. This observation implies a profound underlying similarity among physical systems at criticality, regardless of many aspects of their distinctive microscopic nature. These ideas have found formal expression in the so-called 'universality hypothesis' which, since its inception some 35 years ago, has enjoyed considerable success.

In this course, principal aspects of the contemporary theoretical viewpoint of equilibrium critical phenomena will be reviewed. The ideas of power laws, critical exponents and their relationship to scaling phenomena will be described and set within the context of the powerful renormalisation group technique. The notion of universality as a phenomenological hypothesis will be introduced and its implications for real and model systems will be explored. Finally, the utility of finite-size scaling methods for computer studies of critical phenomena will be discussed, culminating in the introduction of a specific technique suitable for exposing universality in model systems.

2 Background concepts

In seeking to describe near-critical phenomena, it is useful to have a quantitative measure of the difference between the phases coalescing at the critical point: this is the role of the *order parameter*, Q. In the case of the fluid, the order parameter is taken as the difference between the densities of the liquid and vapour phases. In the ferromagnet it is taken as the magnetisation. As its name suggest, the order parameter serves as a measure of the kind of orderliness that sets in when the temperature is cooled below a critical temperature.

Our first task is to give some feeling for the principles which underlie the ordering process. The probability p_a that a physical system at temperature T will have a particular microscopic arrangement (configuration), labelled a, of energy E_a is

$$p_a = \frac{1}{Z} e^{-E_a/k_B T} \tag{1}$$

The prefactor Z^{-1} is the *partition function*: since the system must always have *some* specific arrangement, the sum of the probabilities p_a must be unity, implying that

$$Z = \sum_a e^{-E_a/k_B T} \tag{2}$$

where the sum extends over all possible microscopic arrangements.

In utilising these equations, it is generally correct to suppose that the physical system evolves rapidly (on the timescale of typical observations) amongst all its allowed arrangements, sampling them with the probabilities prescribed by eq. 1; the expectation value of any physical observable O will thus be given by averaging O over all the arrangements a, weighting each contribution by the appropriate probability:

$$\bar{O} = \frac{1}{Z} \sum_a O_a e^{-E_a/k_B T} \tag{3}$$

Sums like eq. 3 are not easily evaluated. Nevertheless, some important insights follow painlessly. Consider the case where the observable of interest is the order parameter, or more specifically the magnetisation of a ferromagnet.

$$Q = \frac{1}{Z} \sum_a Q_a e^{-E_a/k_B T} \tag{4}$$

It is clear from eq. 1 that at very low temperature the system will be overwhelmingly likely to be found in its minimum energy arrangements (ground states). For the ferromagnet, these are the fully ordered spin arrangements having magnetisation $+1$, or -1.

Now consider the high temperature limit. The enhanced weight that the fully ordered arrangement carries in the sum of eq. 4 by virtue of its low energy, is now no longer sufficient to offset the fact that arrangements in which Q_a has some intermediate value, though each carry a smaller weight, are vastly greater in number. A little thought shows that the arrangements which have essentially zero magnetisation (equal populations of up and down spins) are by far the most numerous. At high temperature, these disordered arrangements dominate the sum in eq. 4 and the order parameter is zero.

The competition between energy-of-arrangements weighting (or simply 'energy') and the 'number of arrangements' weighting (or 'entropy') is then the key principle at work here. The distinctive feature of a system with a critical point is that, in the course of this competition, the system is forced to choose amongst a number of macroscopically different sets of microscopic arrangements.

Finally in this section, we note that the probabilistic (statistical mechanics) approach to thermal systems outlined above is completely compatible with classical thermodynamics [1]. Specifically, the bridge between the two disciplines is provided by the following equation

$$F = -k_B T \ln Z \tag{5}$$

where F is the "Helmholtz free energy". All thermodynamic observables, for example the order parameter Q, and response functions such as the specific heat or magnetic susceptibility are obtainable as appropriate derivatives of the free energy. For instance, utilising eq. 2, one can readily verify (try it as an exercise!) that the average internal energy is given by

$$\bar{E} = -\frac{\partial \ln Z}{\partial \beta}, \tag{6}$$

[1] Those of you who have already done PH30029, will have covered this already.

where $\beta = (k_B T)^{-1}$.
The relationship between other thermodynamic quantities and derivatives of the free energy are given in fig. 3

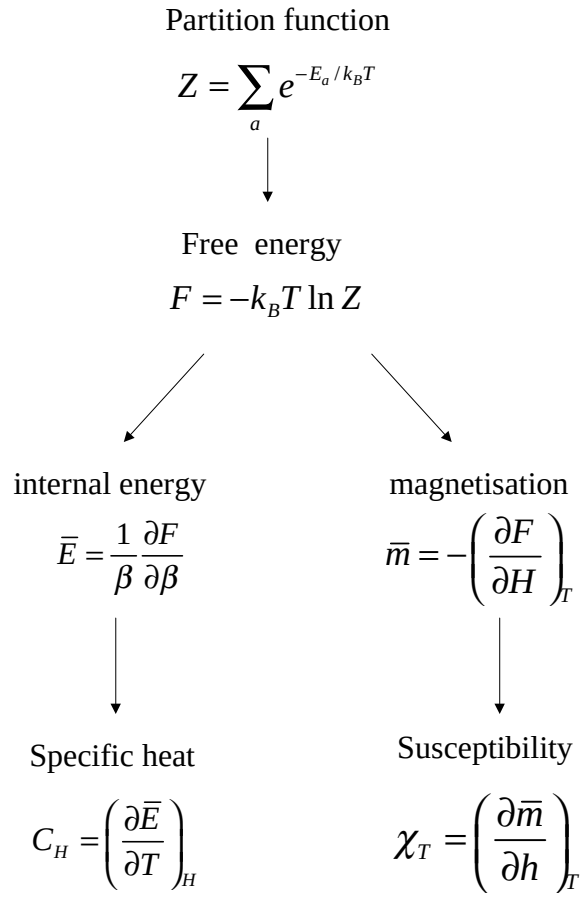


Figure 3: Relationships between the partition function and thermodynamic observables.

3 The Approach to Criticality

It is a matter of experimental fact that the approach to criticality in a given system is characterised by the divergence of various thermodynamic observables. Let us remain with the archetypal example of a critical system, the ferromagnet, whose critical temperature will be denoted as T_c. For temperatures close to T_c, the magnetic response functions (the magnetic susceptibility χ and the specific heat) are found to be singular functions, diverging as the *power* γ of the reduced (dimensionless) temperature $t \equiv (T - T_c)/T_c$:-

$$\chi \equiv \frac{\partial m}{\partial h} \propto t^{-\gamma} \qquad (H = 0) \tag{7}$$

$$C_H \equiv \frac{\partial E}{\partial T} \propto t^{-\alpha} \qquad (H = \text{constant}) \tag{8}$$

Another key quantity is the correlation length ξ, which measures the distance over which fluctuations of the magnetic moments are correlated. This is observed to diverge near the critical point with an exponent ν.

$$\xi \propto t^{-\nu} \qquad (T > T_c,\ H = 0) \tag{9}$$

Similar power law behaviour is found for the order parameter Q (in this case the magnetisation) which vanishes in a singular fashion (it has infinite gradient) as the critical point is is approached as a function of temperature:

$$m \propto t^{\beta} \qquad (T < T_c,\ H = 0), \tag{10}$$

and as a function of magnetic field:

$$m \propto h^{1/\delta} \qquad (T = T_c,\ H > 0). \tag{11}$$

with $h = (H - H_c)/H_c$, the reduced magnetic field.

As examples, the behaviour of the magnetisation and correlation length are plotted in fig. 4 as a function of t.

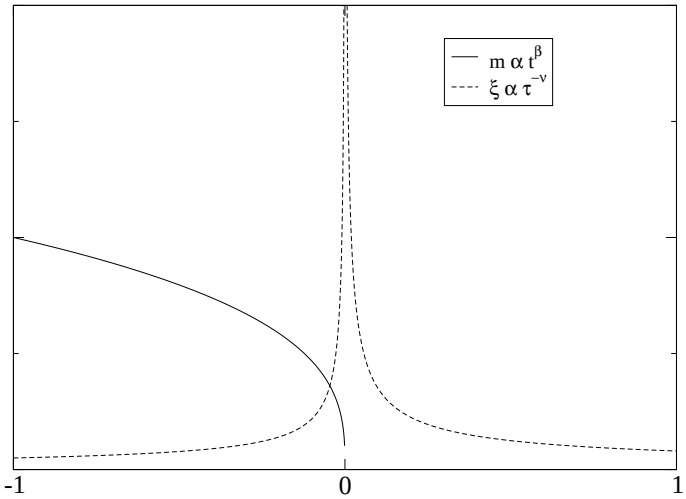


Figure 4: Singular behaviour of the correlation length and order parameter in the vicinity of the critical point.

The quantities $\gamma, \alpha, \nu, \beta$ in the above equations are known as critical exponents. They serve to control the rate at which the various thermodynamic quantities change on the approach to criticality.

Remarkably, the form of singular behaviour observed at criticality for the example ferromagnet also occurs in qualitatively quite different systems such as the fluid. All that is required to obtain the corresponding power law relationships for the fluid is to substitute the analogous thermodynamic quantities in to the above equations. Accordingly the magnetisation order parameter is replaced by the density difference $\rho_{liq} - \rho_{gas}$ while the susceptibility is replaced by the isothermal compressibility and the specific heat capacity at constant field is replaced by the specific heat capacity at constant volume. The approach to criticality in a variety of qualitatively quite different systems can therefore be expressed in terms of a set of critical exponents describing the power law behaviour for that system (see the book by Yeomans for examples).

Even more remarkable is the experimental observation that the values of the critical exponents for a whole range of fluids and magnets (and indeed many other systems with critical points)

are *identical*. This is the phenomenon of *universality*. It implies a deep underlying physical similarity between ostensibly disparate critical systems. The principal aim of theories of critical point phenomena is to provide a sound theoretical basis for the existence of power law behaviour, the factors governing the observed values of critical exponents and the universality phenomenon. Ultimately this basis is provided by the Renormalisation Group (RG) theory, for which K.G. Wilson was awarded the Nobel Prize in Physics in 1982. Historically, however, the first step towards these goals was taken by the static scaling hypothesis.

4 The Static Scaling Hypothesis

The static scaling hypothesis is essentially a plausible conjecture which appears to be consistent with observed phenomena. Its basic assertion is that the singular dependence of the response functions enters through a single variable, namely the reduced temperature (or alternatively the correlation length ξ) and that any other dependence on temperature is smooth and can be regarded as constant over a small temperature range around T_c.

The basis for scaling phenomena in near-critical systems is expressed in the claim that: in the neighbourhood of a critical point, the basic thermodynamic functions (most notably the Free energy) are *generalised homogeneous functions* of their variables. For such functions one can always deduce a scaling law such that by an appropriate change of scale, the dependence on two variables (e.g. the temperature and applied field) can be reduced to dependence on one new variable. This claim may be warranted by the following general argument.

A function of two variables $g(u, v)$ is called a generalised homogeneous function if it has the property

$$g(\lambda^a u, \lambda^b v) = \lambda g(u, v) \tag{12}$$

for all λ, where the parameters a and b (known as scaling parameters) are constants.

Now, the arbitrary scale factor λ can be redefined without loss of generality as $\lambda^a = u^{-1}$ giving

$$g(u, v) = u^{1/a} g(1, \frac{v}{u^{b/a}}) \tag{13}$$

A corresponding relation is obtained by choosing the rescaling to be $\lambda^b = v^{-1}$.

$$g(u, v) = v^{1/b} g(\frac{u}{v^{a/b}}, 1) \tag{14}$$

This equation demonstrates that $g(u, v)$ indeed satisfies a simple power law in *one* variable, subject to the constraint that $u/v^{a/b}$ is a constant. It should be stressed, however, that such a scaling relation specifies neither the function f nor the parameters a and b.

Now, the static scaling hypothesis asserts that in the critical region, the free energy F is a generalised homogeneous function of the thermodynamic fields t and h. Remaining with the example ferromagnet, the following *ad hoc* scaling assumption can then be made:

$$F(\lambda^a t, \lambda^b h) = \lambda F(t, h) \tag{15}$$

Without loss of generality, we can set $\lambda^a = t^{-1}$, implying $\lambda = t^{-1/a}$ and $\lambda^a = t^{-b/a}$.

Then

$$F(t, h) = t^{1/a} F(1, t^{-b/a} h) \tag{16}$$

where our choice of λ ensures that the rhs is now a function of a single variable $t^{-b/a} h$.

Now, as stated in sec. 2 the free energy provides the route to all thermodynamic functions of interest. An expression for the magnetisation can be obtained simply by taking the field derivative of F (cf. fig. 3)

$$m(t,h) = -t^{(1-b)/a} m(1, t^{-b/a} h) \tag{17}$$

In zero applied field $h = 0$, this reduces to

$$m(t,0) = (-t)^{(1-b)/a} m(1,0) \tag{18}$$

where the r.h.s. is a power law in t. Equation (10) then allows identification of the exponent β in terms of the scaling parameters a and b.

$$\beta = \frac{1-b}{a} \tag{19}$$

By taking further appropriate derivatives of the free energy, other relations between scaling parameters and critical exponents may be deduced. Such calculations yield the results $\delta = b/(1-b)$,$\gamma = (2b-1)/a$, and $\alpha = (2a-1)/a$ [2]. Relationships between the critical exponents themselves can be obtained trivially by eliminating the scaling parameters from these equations. The principal results (known as "scaling laws") are:-

$$\alpha + \beta(\delta + 1) = 2 \tag{20}$$

$$\alpha + 2\beta + \gamma = 2 \tag{21}$$

Thus, provided all critical exponents can be expressed in terms of the scaling parameters a and b, then only two critical exponents need be specified, for all others to be deduced. Of course these scaling laws are also expected to hold for the appropriate thermodynamic functions of analogous systems such as the liquid-gas critical point.

4.1 Experimental Verification of Scaling

The validity of the scaling hypothesis finds startling verification in experiment. To facilitate contact with experimental data for real systems, consider again eq. 17. Eliminating the scaling parameters a and b in favour of the exponents β and δ gives

$$\frac{m(t,h)}{t^\beta} = m(1, \frac{h}{t^{\beta\delta}}) \tag{22}$$

where the RHS of this last equation can be regarded as a function of the single scaled variable $\tilde{H} \equiv t^{-\beta\delta} h(t, M)$.

For some particular magnetic system, one can perform an experiment in which one measures m vs h for various fixed temperatures. This allows one to draw a set of isotherms, i.e. $m - h$ curves of constant t. These can be used to demonstrate scaling by plotting the data against the scaling variables $M = t^{-\beta} m(t,h)$ and $\tilde{H} = t^{-\beta\delta} h(t, M)$. Under this scale transformation, it is found that all isotherms (for t close to zero) coincide to within experimental error. Reassuringly, similar results are found using the scaled equation of state of simple fluid systems such as He^3 or Xe.

In summary, the static scaling hypothesis is remarkably successful in providing a foundation for the observation of power laws and scaling phenomena. However, it furnishes little or no

[2] Exercise: try to derive these relationships

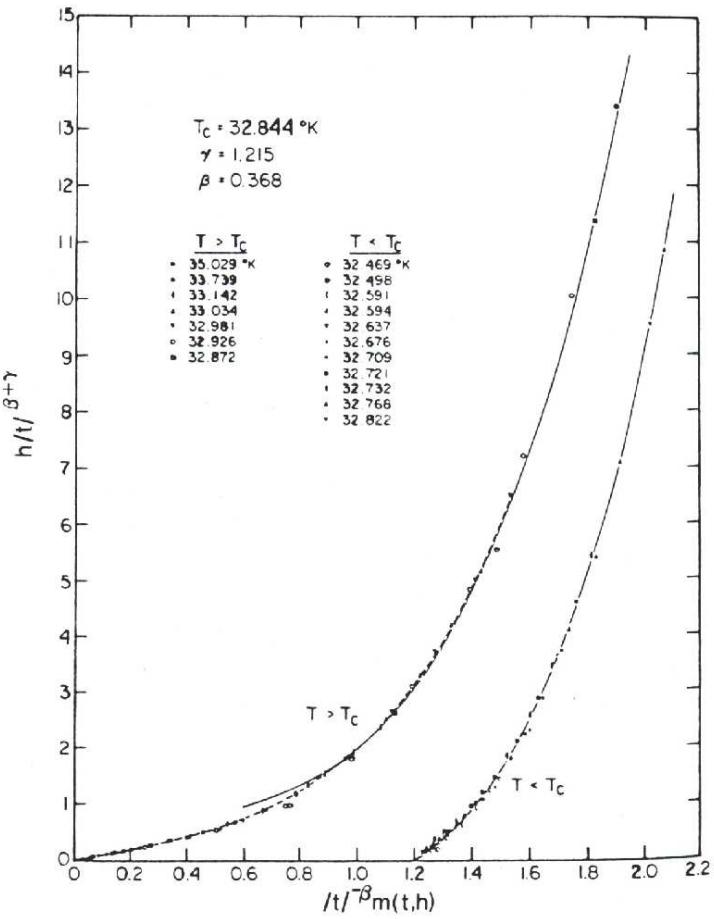


Figure 5: Magnetisation of $CrBr_3$ in the critical region plotted in scaled form (see text).

guidance regarding the role of co-operative phenomena at the critical point. In particular it provides no means for calculating the values of the critical exponents appropriate to given model systems.

5 The Ising model: the prototype model for a phase transition

In order to probe the properties of the critical region, it is common to appeal to simplified model systems whose behaviour parallels that of real materials. The sophistication of any particular model depends on the properties of the system it is supposed to represent. The simplest model to exhibit critical phenomena is the two-dimensional Ising model of a ferromagnet. Actual physical realisations of 2-d magnetic systems do exist in the form of layered ferromagnets such as K_2CoF_4, so the 2-d Ising model is of more than just technical relevance.

The 2-d spin-$\frac{1}{2}$ Ising model envisages a regular arrangement of magnetic moments or 'spins' on an infinite plane. Each spin can take two values, $+1$ ('up' spins) or -1 ('down' spins) and is assumed to interact with its nearest neighbours according to the Hamiltonian

$$\mathcal{H}_I = -J \sum_{<ij>} s_i s_j + H \sum_i s_i \tag{23}$$

where $J > 0$ measures the strength of the coupling between spins and the sum extends over

nearest neighbour spins s_i and s_j. H is a magnetic field term which can be positive or negative (although for the time being we will set it equal to zero). The order parameter is simply the average magnetisation:

$$m = \frac{1}{N}\langle \sum_i s_i \rangle \tag{24}$$

The fact that the Ising model displays a phase transition was argued in sec. 2. At low temperatures for which there is little thermal disorder, there is a preponderance of aligned spins and hence a net spontaneous magnetic moment. As the temperature is raised, thermal disorder increases until at a certain temperature T_c, entropy drives the system through a continuous phase transition to a disordered spin arrangement with zero net magnetisation. These trends are visible in configurational snapshots from computer simulations of the 2D Ising model (see fig. 5). Although each spin interacts only with its nearest neighbours, the phase transition occurs due to cooperative effects among a large number of spins. In the neighbourhood of the transition temperature these cooperative effects engender fluctuations that can extend over all length-scales from the lattice spacing up to the correlation length.

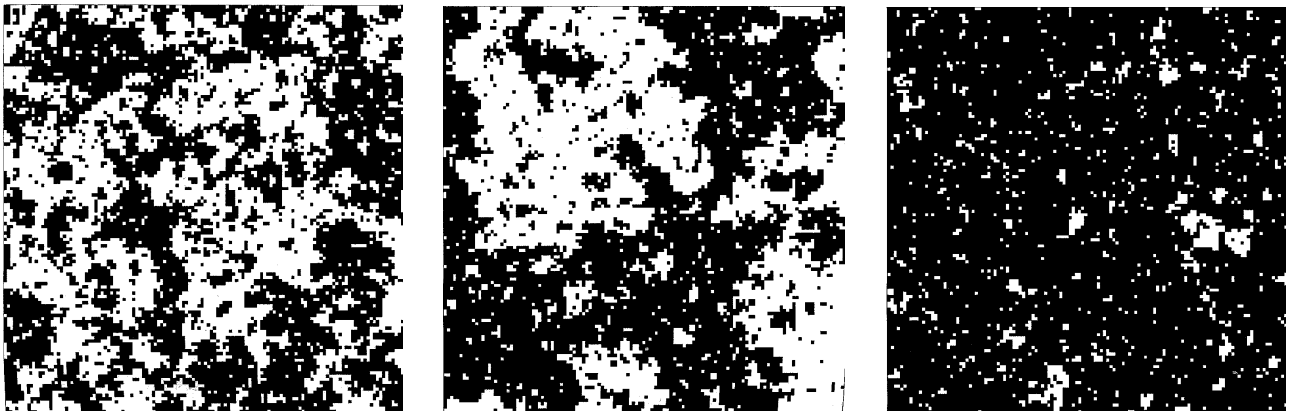


Figure 6: Configurations of the 2d Ising model. The patterns depict typical arrangements of the spins (white=+1, black=−1) generated in a computer simulation of the Ising model on a square lattice of $N = 512^2$ sites, at temperatures of (a) $T = 1.2T_c$, (b) $T = T_c$ and (c) $T = 0.95T_c$. In each case only a portion of the system containing 128^2 sites in shown. The typical island size is a measure of the correlation length ξ: the excess of black over white (below T_c) is a measure of the order parameter.

An interactive Java based Monte Carlo simulation of the Ising model can be found at http://bartok.ucsc.edu/peter/java/ising/ising.html. By altering the temperature you will be able to observe for yourself how the spin arrangements change as one traverses the critical region. Pay particular attention to the configurations near the critical point. They have very interesting properties. We will return to them later!

Although the 2-d Ising model may appear at first sight to be an excessively simplistic portrayal of a real magnetic system, critical point universality implies that many physical observables such as critical exponents are not materially influenced by the actual nature of the microscopic interactions. The Ising model therefore provides a simple, yet *quantitatively* accurate representation of the critical properties of a whole range of real magnetic (and indeed fluid) systems. This universal feature of the model is largely responsible for its ubiquity in the field of critical phenomena.

6 Mean field theory and Perturbation schemes

Of the wide variety of models of interest to the critical point theorist, the majority have shown themselves intractable to direct analytic (pen and paper) assault. In a very limited number of instances models have been solved exactly, yielding the critical exponents and the transition temperature. The 2-d spin-$\frac{1}{2}$ Ising model is certainly the most celebrated such example, its principal critical exponents are found to be $\beta = \frac{1}{8}, \nu = 1, \gamma = \frac{7}{4}$. Its critical temperature is $-2J/\ln(\sqrt{2}-1) \approx 2.269J$. Unfortunately such solutions rarely afford deep insight to the general framework of criticality although they do act as an invaluable test-bed for new and existing theories.

The inability to solve many models exactly often means that one must resort to approximations. One such approximation scheme is mean field theory.

6.1 Mean field solution of the Ising model

Suppose that at some temperature, the order parameter is m. Consider the terms in eq. 23 which contain a particular spin s_0. These terms are, with j restricted to nearest neighbour sites of site 0

$$\begin{aligned} \mathcal{H}(s_0) &= -s_0 \left(J \sum_j s_j + H \right) \\ &= -s_0 \left(qJm + H \right) - J s_0 \sum_j (s_j - m) \end{aligned} \tag{25}$$

where q is the number of nearest neighbours of site 0 ($q = 4$ in 2d, $q = 6$ in 3d). If we disregard the second term in eq. 25, we are left with a non-interacting system. Effectively we are making the approximation that each spin feels a constant effective interaction with its neighbours which is given by the mean of the spin value of each, rather than by their actual instantaneous values. All spins in the system are assumed to feel this 'mean field' interaction.

Let us now consider the statistical mechanics of the system within the mean field approximation. Recalling eq. 2, we can calculate the partition function for spin s_0. Since s_0 can take values $+1$ or -1,

$$Z(s_0) = e^{\beta(qJm+H)} + e^{-\beta(qJm+H)} = 2\cosh(\beta(qJm + H)) \tag{26}$$

Now, since all spins are identical and are assumed non-interacting, the partition function of the whole system is simply

$$Z = [Z(s_0)]^N \tag{27}$$

Thus we find for the free energy (using eq. 5)

$$F = -Nk_BT \ln(2\cosh(\beta(qJm + H))) \tag{28}$$

From which, the magnetisation follows (cf. fig 3) as

$$m = \tanh(\beta(qJm + H)) \tag{29}$$

To find $m(H, T)$, we must numerically solve eq. 29 self consistently.

6.1.1 Zero H solution

In zero field

$$m = \tanh(\frac{mT_c}{T}) \tag{30}$$

where $T_c = qJ/k_B$ is the critical temperature at which m first goes to zero.
We look for a solution where m is small ($\ll 1$). Expanding the tanh function and replacing $\beta = k_B T$ yields

$$m = \frac{mT_c}{T} - \frac{1}{3}\left(\frac{mT_c}{T}\right)^3 + O(m^5) \tag{31}$$

. Then $m = 0$ is one solution. The other solution is given by

$$m^2 = 3\left(\frac{T}{T_c}\right)^3\left(\frac{T_c}{T} - 1\right) \tag{32}$$

Now, considering temperatures close to T_c to guarantee small m, and employing the reduced temperature $t = (T - T_c)/T_c$, one finds

$$m^2 \simeq -3t \tag{33}$$

Hence

$$\begin{array}{lll} m = & 0 & \text{for } T > T_c \text{ since otherwise } m \text{ imaginary} \\ m = & \pm\sqrt{-3t} & \text{for } T < T_c \text{ real} \end{array}$$

6.1.2 Finite (but small) field solution

In a finite, but small field we can make the expansion

$$m = \frac{mT_c}{T} - \frac{1}{3}\left(\frac{mT_c}{T}\right)^3 + \frac{H}{kT} \tag{34}$$

Consider now the isothermal susceptibility

$$\begin{aligned} \chi &\equiv \left(\frac{\partial m}{\partial H}\right)_T \\ &= \frac{T_c}{T}\chi - \left(\frac{T_c}{T}\right)^3 \chi m^2 + \frac{1}{k_B T} \end{aligned}$$

Then

$$\chi\left[1 - \frac{T_c}{T} + \left(\frac{T_c}{T}\right)^3 m^2\right] = \frac{1}{k_B T} \tag{35}$$

Hence near T_c

$$\chi = \frac{1}{k_B T_c}\left(\frac{1}{t + m^2}\right) \tag{36}$$

Then

$$\begin{aligned} \chi = &\quad (k_B T_c t)^{-1} \quad \text{for} \quad T > T_c \\ \chi = &\quad (-2k_B T_c t)^{-1} \quad \text{for} \quad T \leq T_c \end{aligned}$$

where one has to take the non-zero value for m below T_c to ensure +ve χ, i.e. thermodynamic stability.

The schematic behaviour of the Ising order parameter and susceptibility are shown in fig. 6.1.2.

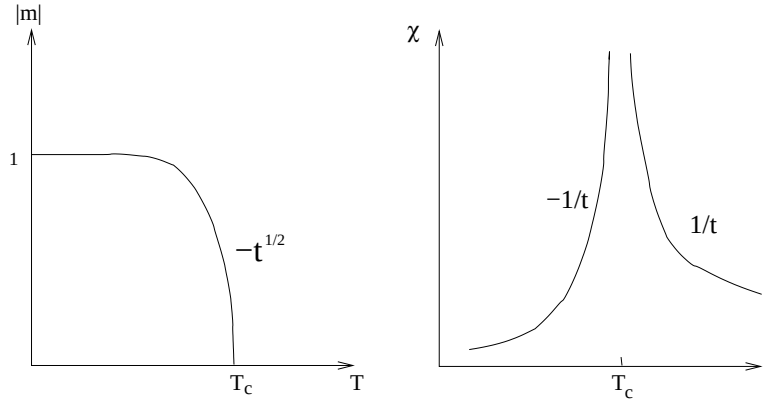


Figure 7: Mean field behaviour of the Ising magnetisation and susceptibility (schematic)

In real ferromagnets, as well as in more sophisticated theories, the exponents β and γ are not the simple fraction and integers found here. This failure of mean field theory to predict the correct exponents is of course traceable to their neglect of correlations.

	Mean Field	$d = 1$	$d = 2$	$d = 3$
Critical temperature $k_B T/qJ$	1	0	0.5673	0.75
Order parameter exponent β	$\frac{1}{2}$	-	$\frac{1}{8}$	0.325 ± 0.001
Susceptibility exponent γ	1	∞	$\frac{7}{4}$	1.24 ± 0.001
Correlation length exponent ν	$\frac{1}{2}$	∞	1	0.63 ± 0.001

Table 1: Comparison of true Ising critical exponents with their mean field theory predictions in a number of dimensions.

6.2 Landau theory

Landau theory is a slightly more general type of mean field theory than that discussed in the previous subsection because it is not based on a particular microscopic model. Its starting point is the Helmholtz free energy, which Landau asserted can be written in terms of a truncated power series expansion of the order parameter. For a ferromagnet this takes the form

$$F(m) = F_0 + a_2 m^2 + a_4 m^4 \tag{37}$$

Here only the terms compatible with the order parameter symmetry are included in the expansion. On symmetry grounds, the free energy of a ferromagnet should be invariant under a reversal of the sign of the magnetisation. Terms linear and cubic in m are not invariant under $m \to -m$, and so do not feature.

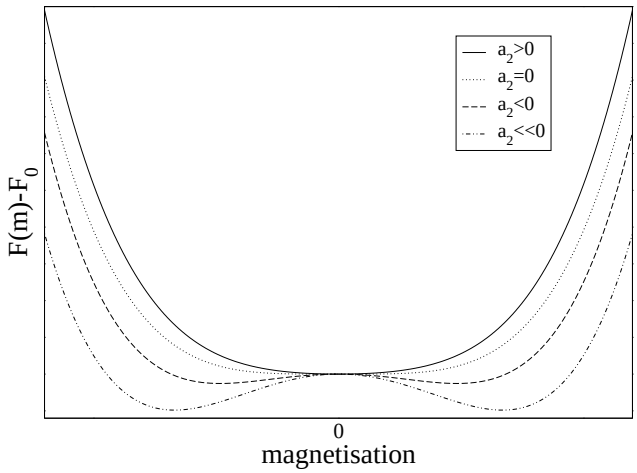


Figure 8: Form of the Landau Free energy for a selection of values of the a_2 parameter.

One can understand how the Landau free energy can give rise to a critical point by plotting its form for various values of a_2 with a_4 assumed positive (which ensures that the magnetisation remains bounded). This is shown in fig. 6.2

Now thermodynamics tells us that the system adopts the state of lowest free energy. Thus from fig. 6.2, we see that for $a_2 > 0$, the system will have $m = 0$, i.e. will be in the disordered (or paramagnetic) phase. For $a_2 < 0$, the minimum in the free energy occurs at a finite value of m, indicating that the ordered (ferromagnetic) phase is the stable one. In fact, the physical (up-down) spin symmetry built into F indicates that there are two equivalent stable states at $m = \pm m_0$. $a_2 = 0$ corresponds to the critical point which marks the border between the ordered and disordered phases.

Clearly a_2 controls the deviation from the critical temperature, and accordingly we may write

$$a_2 = \tilde{a_2} t \tag{38}$$

where t is the reduced temperature.

We can now attempt to calculate critical exponents. Restricting ourselves to the magnetisation exponent β, we first find the equilibrium magnetisation, corresponding to the minimum of the Landau free energy:

$$\frac{dF}{dm} = 2\tilde{a_2} t m + 4a_4 m^3 = 0 \tag{39}$$

which implies

$$m \propto (-t)^{1/2}, \tag{40}$$

so $\beta = 1/2$, which is again a mean field result, as are the other exponents one finds from Landau theory.

6.3 Series expansion

For many years prior to the introduction of the renormalisation group method, the bedrock of information about true (i.e. non mean-field) criticality in model systems was provided by series expansion techniques. These methods seek to deduce results for the critical region using known results obtained away from criticality. One simple example is the high-temperature series which expresses the Boltzmann factor in terms of a temperature expansion with $1/T$ assumed small.

$$\exp(-\mathcal{H}/k_BT) = 1 - \mathcal{H}/k_BT + \frac{1}{2!}(\mathcal{H}/k_BT)^2 + ... \tag{41}$$

Successively higher terms in this series can be regarded as characterising correlations over successively larger distances. Known results for the high-temperature fully disordered phase can be applied to permit the (rather laborious) calculation of the partition function Z using the significant terms in this expansion. This procedure can yield surprisingly accurate numerical results although it is typical that the expansions break down close to the critical point, a failure that can be traced back to the divergent correlation length.

6.4 Computer simulation

In seeking to employ simulation to obtain estimates of bulk critical point properties (such as the location of a critical point and the values of its associated exponents), one is immediately confronted with a difficulty. The problem is that simulations are necessarily restricted to dealing with systems of finite-size and cannot therefore accommodate the truly long ranged fluctuations that characterize the near-critical regime. As a consequence, the critical singularities in C_v, order parameter, etc. appear rounded and shifted in a simulation study. Fig. 9 shows a schematic example for the susceptibility of a magnet.

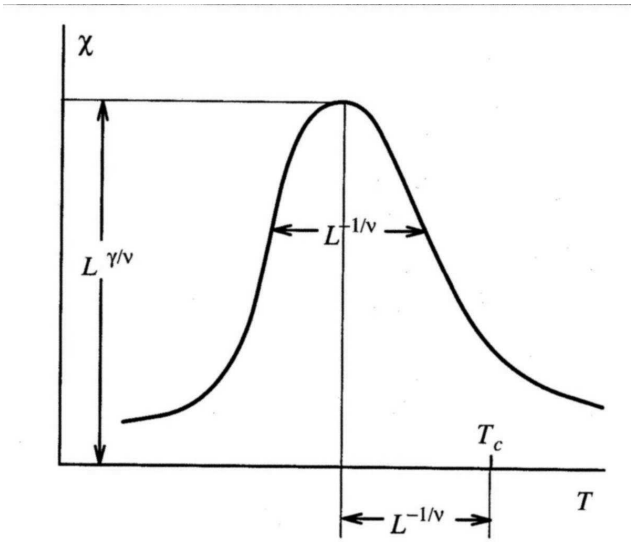


Figure 9: Schematic of the near-critical temperature dependence of the magnet susceptibility in a finite-sized system.

Thus the position of the peak in a response function (such as C_v) measured for a finite-sized system does not provide an accurate estimate of the critical temperature. Although the degree of rounding and shifting reduces with system size, it is often the case, that one computational constraints prevent access to the largest system sizes which would provide accurate estaimates of critical parameters. To help deal with this difficulty, finite-size scaling (FSS) methods have been developed to allow extraction of bulk critical properties from simulations of finite size. FSS will be discussed in 7.7.

7 The Renormalisation Group Theory of Critical Phenomena

The central feature of the critical region is the existence of correlated microstructure on <u>all</u> length-scales up to and including the correlation length. Such a profusion of degrees of freedom can only be accurately characterised by a very large number of variables. As already described, mean field theories and approximation schemes fail in the critical region because they at best incorporate interactions among only a few spins, while neglecting correlations over larger distances. Similarly, the scaling hypothesis fails to provide more than a qualitative insight into the nature of criticality because it focuses on only one length-scale, namely the correlation length itself. Evidently a fuller understanding of the critical region may only be attained by taking account of the existence of structure on all length-scales. Such a scheme is provided by the renormalisation group method, pioneered by K.G. Wilson, which stands today as the cornerstone of the modern theory of critical phenomena.

7.1 The critical point: A many length scale problem

A near critical system can be characterised by three important length scales, namely

(a) The correlation length, ξ, ie the size of correlated microstructure.

(b) Minimum length scale $L_{\min}$, i.e. the smallest length in the microscopics of the problem, e.g. lattice spacing of a magnet or the particle size in a fluid.

(c) Macroscopic size L_{max} eg. size of the system.

The authentic critical region is defined by a window condition:

$$L_{\max} \gg \xi \gg L_{\min} \tag{42}$$

The physics of this regime is hard to tackle by analytic theory because it is characterised by configurational structure on all scales between $L_{\min}$ and ξ (in fact it turns out that the near critical configurational patterns are *fractal*-like, cf. fig. 5(b)). Moreover different length scales are correlated with one another, giving rise to a profusion of coupled variables in any theoretical description. The window regime is also not easily accessed by computer simulation because it entails studying very large system sizes $L_{\max}$, often requiring considerable computing resources.

7.2 Philosophy and Methodology of the RG

The central idea of the renormalisation group (RG) method is a stepwise elimination of the degrees of freedom of the system on successively larger length-scales. To achieve this one introduces a fourth length scale L. In contrast to the other three, which characterise the system itself, L characterises the *description* of the system. It may be thought of as typifying the size of the smallest resolvable detail in a description of the system's microstructure.

Consider the Ising model arrangements displayed in fig. 5(a)-(c). These pictures contain *all* the details of each configuration shown: the resolution length L in this case has its smallest possible value, coinciding with the lattice spacing i.e. $L = L_{\min}$. In the present context, the most detailed description is not the most useful: the essential signals with which we are concerned are hidden in a noise of relevant detail. A clue to eliminating this noise lies in the nature of the correlation length, i.e. the size of the largest droplets. The explicit form of the small scale microstructure is irrelevant to the behaviour of ξ. The small scale microstructure is the noise. To eliminate it, we simply select a larger value of the resolution length (or 'coarse-graining' length) L.

There are many ways of implementing this coarse-graining procedure. We adopt a simple strategy in which we divide our sample into blocks of side L, each of which contains L^d sites, with d the space dimensions . The centres of the blocks define a lattice of points indexed by $I = 1, 2, ..., N/L^d$. We associate with each block lattice point centre, I, a coarse-grained or block variable $S_I(L)$ defined as the spatial average of the local variables it contains:

$$S_I(L) = L^{-d} \sum_{i}^{I} s_i \tag{43}$$

where the sum extends over the L^d sites in the block I. The set of coarse grained coordinates $\{S(L)\}$ are the basic ingredients of a picture of the system having spatial resolution of order L.

The coarse graining operation is easily implemented on a computer. In so doing one is faced with the fact that while the underlying Ising spins can only take two possible values, the block

variables $S_I(L)$ have $L^d + 1$ possible values. Accordingly in displaying the consequences of the blocking procedure, we need a more elaborate colour convention than that used in fig. 5. We will associate with each block a shade of grey drawn from a spectrum ranging from black to white.

The results of coarse-graining configurations typical of three different temperatures are shown in fig. 10 and fig. 11

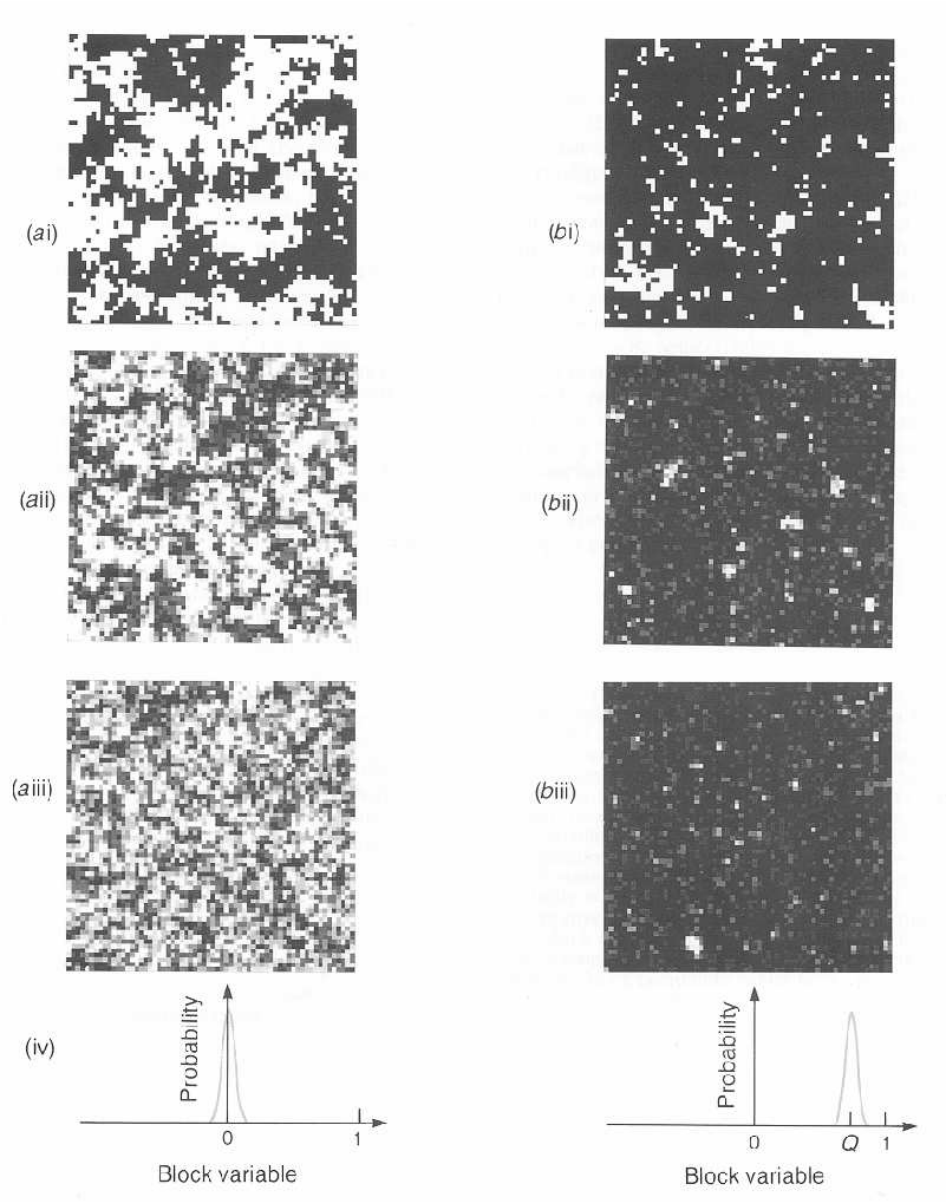


Figure 10: See text for details

Two auxiliary operations are implicit in these results. The first operation is a *length scaling*: the lattice spacing on each blocked lattice has been scaled to the same size as that of the original lattice, making possible the display of correspondingly larger portions of the physical system. The second operation is a *variable scaling*: loosely speaking, we have adjusted the scale ('contrast') of the block variable so as to match the spectrum of block variable values to the spectrum of shades at our disposal.

Consider first a system marginally above its critical point at a temperature T chosen so that the correlation length ξ is approximately 6 lattice spacing units. A typical arrangement (without

coarse-graining) is shown in fig. 10(ai). The succeeding figures, 10(aii) and 10(aiii), show the result of coarse-graining with block sizes $L = 4$ and $L = 8$, respectively. A clear trend is apparent. The coarse-graining *amplifies* the consequences of the small deviation of T from T_c. As L is increased, the ratio of the size of the largest configurational features (ξ) to the size of the smallest (L) is reduced. The ratio ξ/L provides a natural measure of how 'critical' is a configuration. Thus the coarse-graining operation generates a representation of the system that is effectively less critical the larger the coarse-graining length. The limit point of this trend is the effectively fully disordered arrangement shown in fig. 10(aiii) and in an alternative form in fig. 10(aiv), which shows the limiting distribution of the coarse grained variables, averaged over many realisations of the underlying configurations: the distribution is a Gaussian which is narrow (even more so the larger the L value) and centred on zero. This limit is easily understood. When the system is viewed on a scaled L larger than ξ, the correlated microstructure is no longer explicitly apparent; each coarse-grained variable is essentially independent of the others.

A similar trend is apparent below the critical point. Fig. 10(bi) show a typical arrangement at a temperature $T < T_c$ such that again ξ is approximately 6 lattice spacings. Coarse-graining with $L = 4$ and $L = 8$ again generates representations which are effectively less critical (figs. 10(bii) and (biii)). This time the coarse-graining smoothes out the microstructure which makes the order incomplete. The limit point of this procedure is a homogeneously ordered arrangement in which the block variables have a random (Gaussian) distribution centred on the order parameter (fig. 10(biv)).

Consider now the situation *at* the critical point. Fig. 11(ai) shows a typical arrangement; figs. 11(aii) and (aiii) show the results of coarse-graining with $L = 4$ and $L = 8$ respectively. Since the correlation length is as large as the system itself the coarse graining does not produce less critical representations of the physical system: each of the figures displays structure over *all* length scales between the lower limit set by L and the upper limit set by the size of the display itself. A limiting trend is nevertheless apparent. Although the $L = 4$ pattern is qualitatively quite different from the pattern of the local variables, the $L = 4$ and $L = 8$ patterns display qualitatively similar features. These similarities are more profound than is immediately apparent. A statistical analysis of the spectrum of $L = 4$ configurations (generated as the local variables evolve in time) show that it is almost identical to that of the $L = 8$ configurations (given the block variable scaling). This state of affairs is expressed in fig. 11(iv), which shows the near coincidence of the distribution of block variables (grey-levels) for the two different coarse-graining lengths. The implication of this limiting behaviour is clear: the patterns formed by the ordering variable at criticality look the same (in a statistical sense) when viewed on all sufficiently large length scales.

Let us summarise. Under the coarse-graining operation there is an evolution or *flow* of the system's configuration spectrum. The flow tends to a limit, or fixed point, such that the pattern spectrum does not change under further coarse-graining. These scale-invariant limits have a trivial character for $T > T_c$, (a perfectly disordered arrangement) and $T < T_c$, (a perfectly ordered arrangement). The hallmark of the critical point is the existence of a scale-invariant limit which is neither fully ordered nor fully disordered but which possesses structure on all length scales. The answers we seek are to be found in the nature and, indeed, the very existence of this scale-invariant limit.

7.3 Universality and Scaling

Armed with the coarse-graining technique, we now address the problems posed by the universality phenomenon. We seek to understand how it is that systems as different microscopically, as the fluid and the magnet can nevertheless display critical point behaviour which (in certain

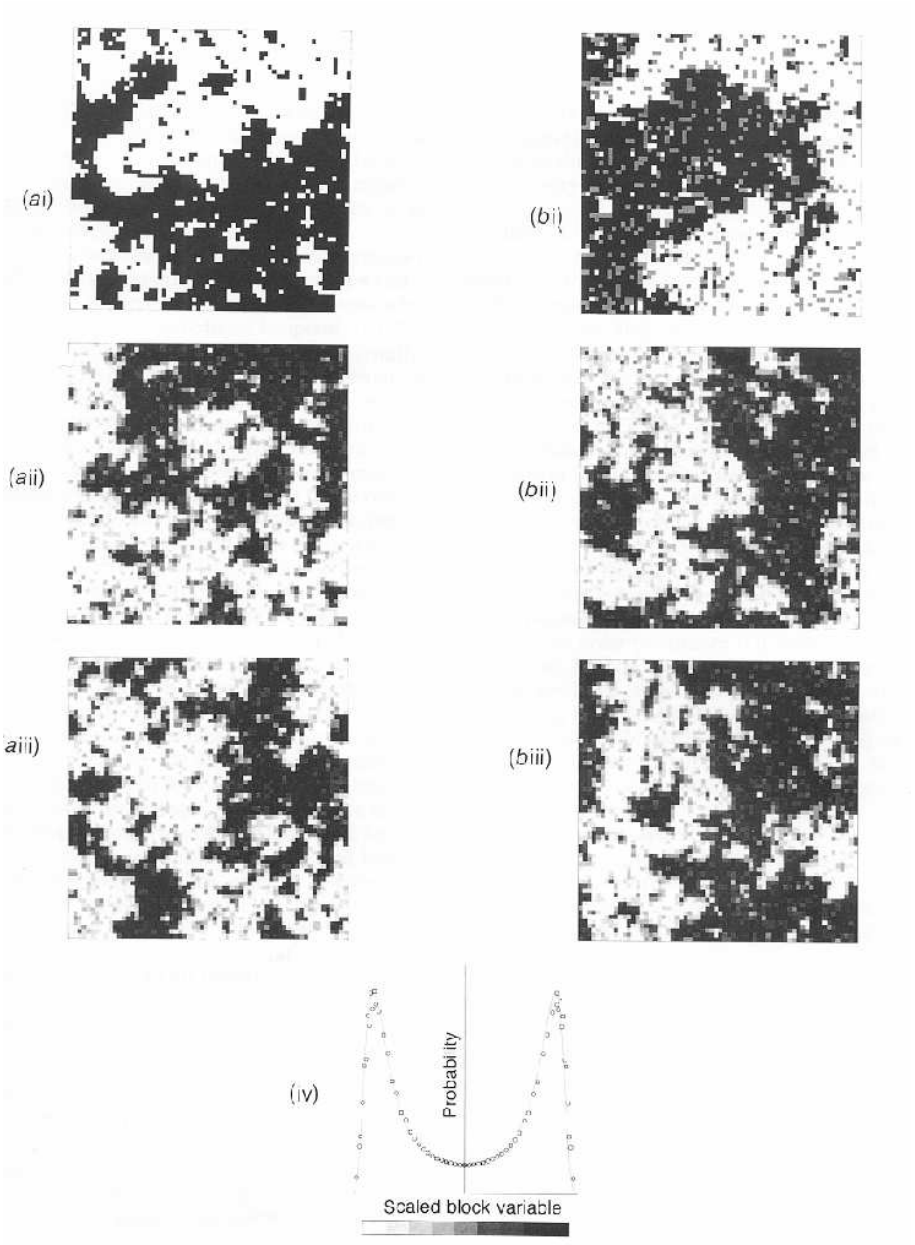


Figure 11: See text for details

respects) is quantitatively identical.

We will postpone an explicit examination of the behaviour of the fluid and the magnet to sec. 7.4. The essential points can be made more readily by a comparative study of two different models. To that end we introduce a variant of the Ising model. In the original model the local spin variables assume just two values (1 and -1). In the new model, the spins take on three values (1, 0 and -1) In other respects the models are the same. To distinguish between them we will refer to them as the two-state and three state Ising models. The two models have properties which are clearly different: for example, T_c for the three-state model is some 30% lower than that of the two-state model (for the same coupling J). However, there is abundant evidence that the two models (like the fluid-magnet pair) have the same universal properties. Let us explore what is the same and what is different in the configurations of the two models.

The configurations of the local variables s_i are clearly qualitatively different for the two models: at this level the two-state character of the one model and the three-state character of the others are quite apparent (fig. 11ai,bi). Now however, consider the coarse-grained configurations (with $L = 4$ and $L = 8$ respectively) for the three-state model at the critical point. We have already seen that the coarse-graining operation bears the configuration spectrum of the critical two-state Ising model to a non-trivial scale-invariant limit. It is scarcely surprising that the same is true for the three-state model. What is remarkable is that the two limits are the same! The coarse-graining erases the physical differences apparent in configurations where the local behaviour is resolvable, and exposes a profound configurational similarity.

Some qualification is necessary here. There is one configurational difference-a difference between the overall *scales* of the block variables of the two problems–which the coarse graining operation itself does not erase. This difference is significant, but superficial. It can be suppressed (and the near-critical perfect statistical similarity of the configurations exposed) merely by means of the variable scaling we introduced as part of the coarse-graining procedure.

The similarity of the configurations of the different systems is not restricted to the critical temperature itself. Suppose we have a two state model and a three state model each somewhat above their critical points at reduced temperature t. The two systems will have somewhat different correlation lengths, ξ_1 and ξ_2 say, manifesting the non-universality of the proportionality constant in eq. 9. This difference is less than meets the eye. Although it survives the the coarse-graining itself, it can be suppressed by the combined effects of the coarse graining and the auxiliary *length-scale* operation. Specifically, we choose coarse-graining lengths L_1 for L_2 for the two models such that $\xi_1/L_1 = \xi_2/L_2$. We adjust the scales of the block variables (our grey level control) so that the typical variable value is the same for the two systems. We adjust the length scale of the systems (stretch or shrink our snapshots) so that the sizes of the minimum-length-scale structure (set by L_1 and L_2) looks the same for each system. One can show that these operations take care of all the differences : the configurations of the two systems again look (statistically) identical. Precisely what they look like depends upon our choice if ξ/L.

7.4 Fluid-magnet universality

The similarities in the critical behaviour of fluids and magnets can be traced to the underlying similarity in their coarse-grained configurations (cf fig. 12).

In a magnet, the relevant configurations are those formed by the coarse-grained magnetisation (the magnetic moment averaged over a block of side L). In a fluid, the relevant configurations are those of the coarse-grained density (the mass averaged over a block if side L) or more precisely, its fluctuation from its macroscopic average. The patterns in the latter (bubbles of liquid or vapour) may be matched to pattern in the former (microdomains of the magnetisation), given appropriate scaling operations to camouflage the differences between the length

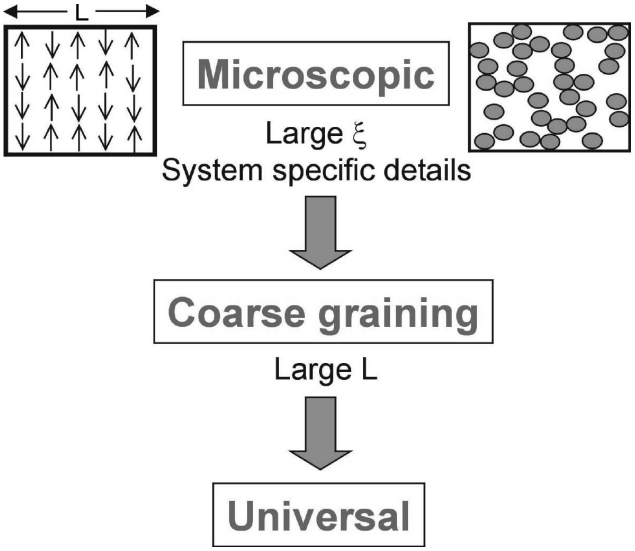


Figure 12: Schematic representation of the coarse graining operation via which the universal properties of fluids and magnets may be exposed.

scales and the differences between the variable scales (indeed, the variable *units*: density and magnetisation) in the two problems (cf. fig. 13).

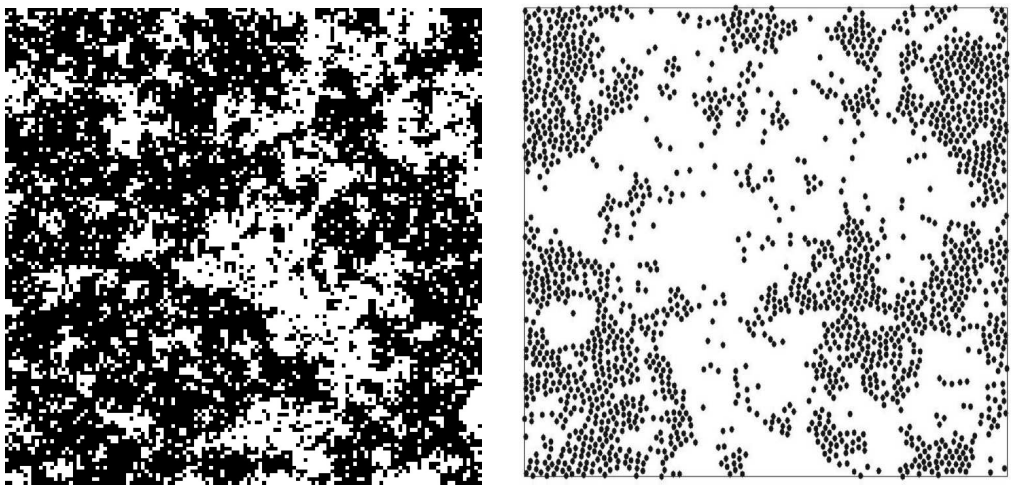


Figure 13: Snapshot configurations of the 2D critical Ising model (left) and the 2D critical Lennard-Jones fluid (right). When viewed on sufficiently large length scales the configurational patterns appear universal and self similar.

7.5 Universality classes

Coarse graining does not erase all differences between the physical properties of critical systems. Differences in the space dimension d of two critical systems will lead to different universal properties such as critical indices. Thus, for instance, the critical exponents of the 2D magnet, match those of the 2d fluid, but they are different to those of 3d magnets and fluids.

In fact the space dimension is one of a small set of qualitative features of a critical system which are sufficiently deep-seated to survive coarse graining and which together serve to define the system's universal behaviour, or *universality class*. The constituents of this set are not all identifiable *a priori*. They include the number of components n of the order parameter. Up

to now, we have only considered order parameters which are scalar (for a fluid the density, for a magnet the magnetisation), for which $n = 1$. In some ferromagnets, the order parameter may have components along two axes, or three axes, implying a vector order parameter, with $n = 2$ or $n = 3$, respectively. It is clear that the order-parameter n-value will be reflected in the nature of the coarse-grained configurations, and thus in the universal observables they imply.

A third important feature which can change the universality class of a critical system is the range of the interaction potential between its constituent particles. Clearly for the Ising model, interactions between spins are inherently nearest neighbour in character. Most fluids interact via dispersion forces (such as the Lennard-Jones potential) which is also short ranged owing to the r^{-6} attractive interaction. However some systems have much longer ranged interactions. Notable here are systems of charged particles which interact via a Coulomb potential. The long ranged nature of the Coulomb potential (which decays like r^{-1}) means that charged systems often do not have the same critical exponents as the Ising model and fluid.

7.6 Critical indices

We now turn to show how the prototype universal quantities, the critical indices, may be computed from the properties of the coarse-grained configuration spectrum. To do so, we must express the findings of the preceding section in a more mathematically explicit form.

Consider then, the coarse-grained variables of our Ising model system, or, indeed, of any system belonging to the same universality class. In what follows we will refer only to the behaviour of a single typical coarse grained variable, which we shall denote $S(L)$. We suppose that the system is sufficiently close to criticality (at sufficiently small t) that the important observables are dominated by large-scale microstructure: that is the critical condition $\xi \gg L_{\min}$. The configurational universality developed above may then be expressed in the claim that, for *any* L and t satisfying the prescribed conditions, scale factors $A(L)$ and $B(L)$ may be found such that the coarse grained probability distribution $p(S(L))$ can be written in the form

$$p(S(L), t) = B(L)\tilde{p}(A(L)t, B(L)S(L)) \tag{44}$$

where $\tilde{p}$ is a function unique to a universality class. The role of the scale factors A and B is to absorb the basic non-universal scales identified in the preceding section. The information we seek is implicit in the manner in which these scale factors depend upon the coarse-graining length L. The key results are

$$\begin{aligned} A(L) &= A_0 L^{1/\nu} \\ B(L) &= B_0 L^{\beta/\nu} \end{aligned} \tag{45}$$

The amplitudes A_0 and B_0 are system specific (non-universal) constants.

The results reveal that the basic critical indices (in the form $1/\nu$ and β/ν) serve to characterise the ways in which the configuration spectrum evolves under coarse-graining. Consider, first the index β/ν. Precisely at the critical point, there is only one way in which the coarse-grained configurations change with L: the overall scale of the coarse-grained variable (the black-white contrast in our grey scale representation) is eroded with increasing L. Thus the configurations of coarse-graining length L_1 match those of a larger coarse-graiing length L_2 only if the variable scale in the latter configurations is amplified. The required amplification follows from eqns. 44 and 45: it is $B(L_2)/B(L_1) = (L_2/L_1)^{\beta/\nu}$. The index β/ν thus controls the rate at which the scale of the ordering variable decays with increasing coarse-graining length.

Consider now the index $1/\nu$. For small but non-zero reduced temperature (large but finite ξ) there is second way in which the configuration spectrum evolves with L. As noted previously,

coarse graining reduces the ratio of correlation length to coarse-graining length, and results in configurations with a less critical appearence. More precisely, we see from eq. 44 that increasing the coarse graining length from L_1 to L_2 while keeping the reduced temperature constant has the same effect on the configuration spectrum as keeping coarse-graining length constant which amplifying the reduced temperature t by a factor $A(L_2)/A(L_1) = (L_2/L_1)^{1/\nu}$. One may think of the combination $A(L)t$ as a measure of the effective reduced temperature of the physical system viewed with resolution length L. The index $1/\nu$ controls the rate at which the effective reduced temperature frows with increasing coarse-graining length.

7.7 Finite-size scaling

Consider the average of the block variable $S(L)$. Consideration of eq. 43 shows that this is non other than the value of the order parameter Q, measured over a block of side L. Now the block variable average $\bar{S}(L)$ is defined by the first moment of the probability distribution p, i.e.

$$Q(L,t) = \bar{S}(L,t) = \int S(L)p(S(L),t)dS(L) \tag{46}$$

Making use of the representation of eq. 44, we then find that

$$\begin{aligned} Q(L,t) &= \int S(L)B(L)\tilde{p}(A(L)t, S(L)B(L))dS(L) \\ &= B^{-1}(L)\int S(L)B(L)\tilde{p}(A(L)t, S(L)B(L))d(S(L)B(L)) \\ &= B^{-1}(L)f(A(L)t) \\ &= B_0 L^{-\beta/\nu} f(A_0 L^{1/\nu} t) \end{aligned} \tag{47}$$

where f is a universal function (defined by the first moment of $\tilde{p}$).

The above results provide us with a prescription for determining the critical exponent ratios β/ν and $1/\nu$ via computer simulations of near critical systems. For instance, at the critical point ($t = 0$) and for finite block size, $Q(L, 0)$ will not be zero (the T at which Q vanishes for finite L is above the true T_c, cf. sec 6.4). However, we know that its value must vanish in the limit of infinite L; it does so like

$$Q(L,0) = B_0 L^{-\beta/\nu} f(0) \equiv Q_0 L^{-\beta/\nu} \tag{48}$$

Thus by studying the critical point L dependence of Q we can estimate β/ν. A similar approach in which we study two block sizes L, and tune t separately in each case so that the results for $QL^{\beta/\nu}$ are identical provides information on the value of $1/\nu$.

8 The renormalisation group: effective coupling viewpoint

Let us begin by returning to our fundamental equation (eq. 1), which we rewrite as

$$p = Z^{-1} e^{-\mathcal{H}} \tag{49}$$

where $\mathcal{H} \equiv E/k_B T$.

The first step is then to imagine that we generate, by a computer simulation procedure for example, a sequence of configurations with relative probability $\exp(-\mathcal{H})$. We next adopt some coarse-graining procedure which produces from these original configurations a set of coarse-grained configurations. We then ask the question: what is the energy function $\mathcal{H}'$ of the coarse-grained variables which would produce these coarse-grained configurations with the correct relative probability $\exp(-\mathcal{H}')$? Clearly the form of $\mathcal{H}'$ depends on the form of $\mathcal{H}$ thus we can write symbolically

$$\mathcal{H}' = R(\mathcal{H}) \tag{50}$$

The operation R, which defines the coarse-grained configurational energy in terms of the microscopic configurational energy function is known as a renormalisation group transformation (RGT). What it does is to replace a hard problem by a less hard problem. Specifically, suppose that our system is near a critical point and that we wish to calculate its large-distance properties. If we address this task by utilising the configurational energy and appealing to the basic machinery of statistical mechanics set out in eqs. 1 and 2, the problem is hard. It is hard because the system has fluctuations on all the (many) length scales intermediate between the correlation length ξ and the minimum length scale $L_{\min}$.

However, the task may instead be addressed by tackling the statistical mechanics of the coarse-grained system described by the energy $\mathcal{H}'$. For the large-distance properties of this coarse-grained system are the same as the large-distance properties of the physical system, since the coarse-graining operation preserves large-scale configurational structure. In this representation the problem is a little easier. For, while the correlation length associated with $\mathcal{H}'$ is the same as the correlation length associated with $\mathcal{H}$, the minimum length scale of $\mathcal{H}'$ is bigger than that of $\mathcal{H}$, by virtue of the coarse- graining operation. Thus the statistical mechanics of $\mathcal{H}'$ poses a not-quite-so-many-length-scale problem, a problem which is effectively a little less critical and is thus a little easier than that posed by the statistical mechanics of $\mathcal{H}$. The benefits accruing from this procedure may be amplified by repeating it. Repeated application of the operation R will eventually result in a coarse- grained energy function describing configurations in which the correlation length is no bigger than the minimum length scale. The associated coarse-grained system is far from criticality. Its properties may be reliably computed by any of a wide variety of approximation schemes available for dealing with noncritical systems. These properties are the desired large-distance properties of the physical system. In effect, repeated application of the RGT captures successively the effects of larger and larger scale fluctuations upon the largest-scale fluctuations of concern; as explicit reference to fluctuations of a given scale is eliminated by coarse-graining, their effects are carried forward implicitly in the parameters of the coarse-grained energy.

In order to put some flesh on this formalism, and to provide a framework for a simple illustrative calculation, let us return to the lattice Ising model explored in the previous sections, for which the energy function depended only on the product of nearest neighbour spins. The coefficient of this product in the energy is the exchange coupling, J. In principle, however, other kinds of interactions are also allowed; for example, we may have a product of second neighbour spins with strength J_2 or, perhaps, a product of four spins (at sites forming a square whose side is

the lattice spacing), with strength J_3. Such interactions in a real magnet have their origin in the quantum mechanics of the atoms and electrons and clearly depend upon the details of the system. For generality therefore we will allow a family of exchange couplings J_1,J_2,$J_3, \ldots$, or $J_a, a = 1, 2, \ldots$ In reduced units, the equivalent coupling strengths are $K_a = J_a/k_BT$. Their values determine uniquely the energy for any given configuration.

Now consider the coarse-graining procedure. Let us suppose that this procedure takes the form of a'majority rule' operation in which the new spins are assigned values $+1$ or -1 according to the signs of the magnetic moments of the blocks with which they are associated. After this coarse-graining procedure, the new energy function $\mathcal{H}'$ will be expressible in terms of some new coupling strengths K' describing the interactions amongst the new spin variables (and thus, in effect, the interactions between blocks of the original spin variables). The RGT simply states that these new couplings depend on the old couplings: K'_1 is some function f_1 of all the original couplings, and generally

$$K'_a = f_a(K_1, K_2, \ldots) = f_a(\mathbf{K}), \;\; a = 1, 2, \ldots \tag{51}$$

where $\mathbf{K}$ is shorthand for the set $K_1, K_2, \ldots$

We note that it is not only useful to allow for arbitrary kinds of interactions: if we wish to repeat the transformation several (indeed many) times, it is also necessary because even if we start with only the nearest neighbour coupling in $\mathcal{H}$ the transformation will in general produce others in $\mathcal{H}'$.

8.1 A simple example

This subsection is included for completeness and does not form part of the examinable course material

This example illustrates how one can perform the renormalisation group transformation (51) directly, without recourse to a 'sequence of typical configurations'. The calculation involves a very crude approximation which has the advantage that it simplifies the subsequent analysis.

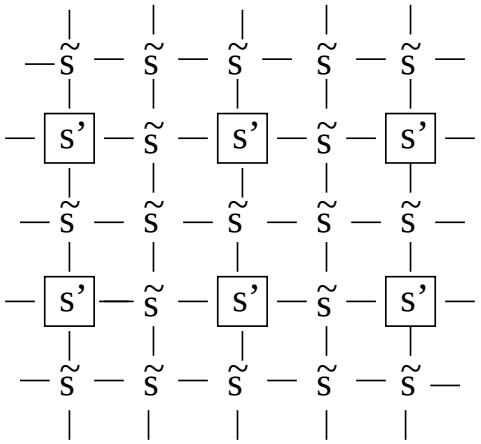


Figure 14: Coarse graining by decimation. The spins on the original lattice are divided into two sets $\{s'\}$ and $\{\tilde{s}\}$. The $\{s'\}$ spins occupy a lattice whose spacing is twice that of the original. The effective coupling interaction between the $\{s'\}$ spins is obtained by performing the configurational average over the $\{\tilde{s}\}$

Consider an Ising lattice model in two dimensions, with only nearest neighbour interactions as shown in fig. 14. We have divided the spins into two sets, the spins $\{s'\}$ form a square lattice

of spacing 2, the others being denoted by $\{\tilde{s}\}$. One then defines an effective energy function $\mathcal{H}'$ for the s' spins by performing an average over all the possible arrangements of the $\tilde{s}$ spins

$$\exp(-\mathcal{H}') = \sum_{\{\tilde{s}\}} \exp(-\mathcal{H}). \tag{52}$$

This particular coarse-graining scheme is called 'decimation' because a certain fraction (not necessarily one-tenth!) of spins on the lattice is eliminated. This formulation of a new energy function realises two basic aims of the renormalisation group method: the long-distance physics of the 'original' system, described by $\mathcal{H}$, is contained in that of the 'new' system, described by $\mathcal{H}'$ (indeed the partition functions are the same) and the new system is further from critically because the ratio of correlation length to lattice spacing ('minimum length scale') has been reduced by a factor of 1/2 (the ratio of the lattice spacings of the two systems). We must now face the question of how to perform the configuration sum in eq. 52. This cannot in general be done exactly, so we must resort to some approximation scheme. The particular approximation which we invoke is the high temperature series expansion which was touched upon in sec. 6.3. In its simplest mathematical form, since $\mathcal{H}$ contains a factor $1/k_BT$, it involves the expansion of $\exp(-\mathcal{H})$ as a power series:

$$\exp(-\mathcal{H}/k_BT) = 1 - \mathcal{H}/k_BT + \frac{1}{2!}(\mathcal{H}/k_BT)^2 + ... \tag{53}$$

We substitute this expansion into the right hand side of eq. 52 and proceed to look for terms which depend on the s' spins after the sum over the possible arrangements of the $\tilde{s}$ spins is performed. This sum extends over all the possible (± 1) values of all the $\tilde{s}$ spins. The first term (the 1) in the expansion of the exponential is clearly independent of the values of the s' spins. The second term ($\mathcal{H}$) is a function of the s' spins, but gives zero when the sum over the s' spins is performed because only a single factor of any s' ever appears, and $+1-1=0$. The third term ($\mathcal{H}^2/2$) does contribute. If one writes out explicitly the form of $\mathcal{H}^2/2$ one finds terms of the form $K^2 s'_1 \tilde{s}\tilde{s} s'_2 = K^2 s'_1 s'_2$, where s'_1 and s'_2 denote two spins at nearest neighbour sites on the lattice of s' spins and $\tilde{s}$ is the spin (in the other set) which lies between them. Now, in the corresponding expansion of the left hand side of eq. 52, we find terms of the form $K' s'_1 s'_2$, where K' is the nearest neighbour coupling for the s' spins. We conclude (with a little more thought than we detail here) that

$$K' = K^2 \tag{54}$$

Of course many other terms and couplings are generated by the higher orders of the high temperature expansion. If our aim is to produce reliable values for the critical temperature and exponents it is necessary to include these terms. However, our aim here is to use this simple calculation to illustrate the renormalisation group method. Let us therefore close our eyes, forget about the higher order terms and show how the RGT (eq. 54) can be used to obtain information on the phase transition. The first point to note is that eq. 54 has the fixed point $K^* = 1$; if $K = 1$ then the new effective coupling K' has the same value 1. Further, if K is just larger than 1, then K' is larger than K, i.e. further away from 1. Similarly, if K is less than 1, K' is less than K. We say that the fixed point is unstable: the flow of couplings under repeated iteration of eq. 54 is away from the fixed point, as illustrated in fig. 15.

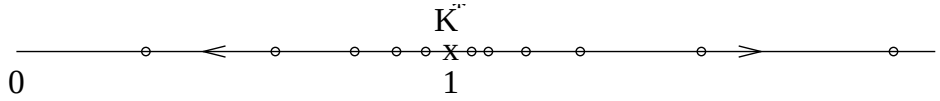


Figure 15: Coupling flow under the decimation transformation described in the text.

To expose the physical significance of these results let us suppose that the original system is at its critical point so that the ratio of correlation length to lattice spacing is infinite. After one application of the decimation transformation, the effective lattice spacing has increased by a factor of two, but this ratio remains infinite; the new system is therefore also at its critical point. Within the approximations inherent in eq. 54, the original system is an Ising model with nearest neighbour coupling K and the new system is an Ising model with nearest neighbour coupling K'. If these two systems are going to be at a common critically, we must identify $K' = K$. The fixed point $K^* = 1$ is therefore a candidate for the critical point K_c, where the phase transition occurs. This interpretation is reinforced by considering the case where the original system is close to, but not at, criticality. Then the correlation length is finite and the new system is further from critically because the ratio of correlation length to lattice spacing is reduced by a factor of two. This instability of a fixed point to deviations of K from K^* is a further necessary condition for its interpretation as a critical point of the system. In summary then we make the prediction

$$K_c = J/k_B T_c = 1 \tag{55}$$

We can obtain further information about the behaviour of the system close to its critical point. In order to do so, we rewrite the transformation (eq. 54) in terms of the deviation of the coupling from its fixed point value. Simple algebra shows that

$$K' - K^* = 2(K - K^*) + (K - K^*)^2 \tag{56}$$

For a system sufficiently close to its critical temperature the final term can be neglected. The deviation of the coupling from its fixed point (critical) value is thus bigger for the new system than it is for the old by a factor of two. This means that the reduced temperature is also bigger by a factor of two:

$$t' = 2t \tag{57}$$

But the correlation length (in units of the appropriate lattice spacing) is smaller by a factor of $1/2$:

$$\xi' = \xi/2 \tag{58}$$

Thus, when we double t, we halve ξ, implying that

$$\xi \propto t^{-1} \tag{59}$$

for T close to T_c. Thus we see that the RGT predicts scaling behaviour with calculable critical exponents. In this simple calculation we estimate the critical exponent $\nu = 1$ for the square lattice Ising model. This prediction is actually in agreement with the exactly established value. The agreement is fortuitous. A more sober measure of the crudity of the calculation is to be found in the fact that the prediction in equation for K_c, is larger than the exactly established value by a factor of more than two. In order to obtain reliable estimates more sophisticated and systematic methods must be used.

The crude approximation in the calculation above produced a transformation, eq. 54, involving only the nearest neighbour coupling, with the subsequent advantages of simple algebra. We pay a penalty for this simplicity in two ways: the results obtained for critical properties are in rather poor agreement with accepted values, and we gain no insight into the origin of universality.

8.2 Universality and scaling

In order to expose how universality can arise, we should from the start allow for several different kinds of coupling J_a, and show how the systems with *different* J_a can have the *same* critical behaviour.

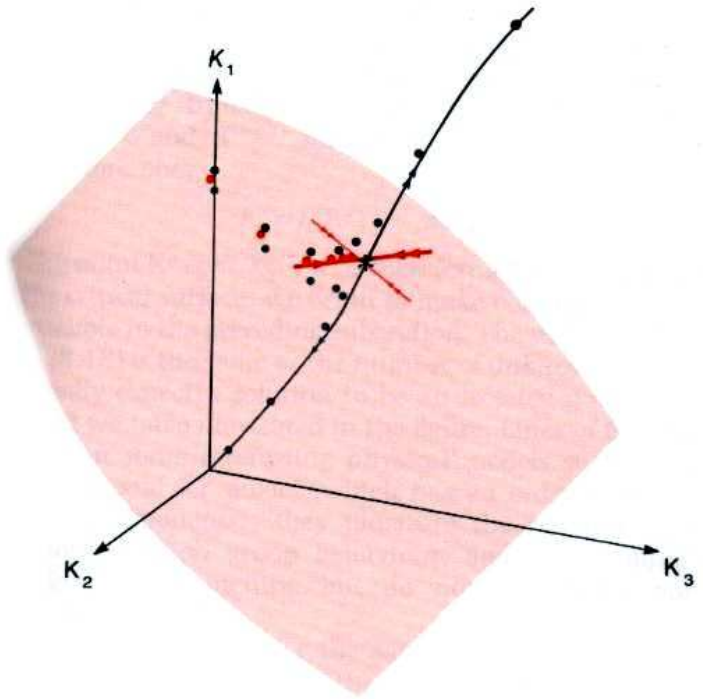


Figure 16: General flow in coupling space

Fig. 16 is a representation of the space of all coupling strengths K_a in the energy function $\mathcal{H}/k_BT$. This is of course actually a space of infinite dimension, but representing three of these, as we have done, enables us to illustrate all the important aspects. First let us be clear what the points in this space represent. Suppose we have some magnetic material which is described by a given set of exchange constants J_1, J_2, J_3..... As the temperature T varies, the coupling strengths $K_a = J_a/k_BT$ trace out a straight line, or ray, from the origin of the space in the direction $(J_1, J_2, J_3$....). Points on this ray close to the origin represent this magnet at high temperatures, and conversely points far from the origin represent the magnet at low temperatures. The critical point of the magnet is represented by a specific point on this ray, $K_a = J_a/k_BT, a = 1, 2, \ldots$ The set of critical points on all of the possible rays forms a *surface*, the critical surface. Formally, it is defined by the set of all possible models (of the Ising type) which have infinite correlation length. It is shown schematically as the shaded surface in fig. 16. (In the figure it is a two-dimensional surface; more generally it has one dimension less than the full coupling constant space, dividing all models into high and low tempera- ture phases.)

Our immediate goal then is to understand how the RGT can explain why different physical systems near this critical surface have the same behaviour. Let us turn now to the schematic representation of the RG flow in fig. 16. Suppose we start with a physical system, with coupling strengths $K_a, a = 1, 2, \ldots$. What the RGT does is generate a new point in the figure, at the coupling strengths $K_a^{(1)} = f_a(\mathbf{K})$; these are the couplings appearing in the effective energy function describing the coarse-grained system. If we repeat the transformation, the new energy function has coupling strengths $K_a^{(2)} = f_a(\mathbf{K})$. Thus repeated application of the transformation generates a flow of points in the figure: $\mathbf{K} \to \mathbf{K^{(1)}} \to \ldots \to \mathbf{K^{(n)}}$ where the superscript (n) labels

the effective couplings after n coarse-graining steps. if the change in coarse-graining scale is b (> 1) at each step, the total change in coarse-graining scale is b^n after n steps. In the process, therefore, the ratio of correlation length to coarse-graining scale is reduced by a factor of b^{-n}. The dots in fig. 16 identify three lines of RG flow starting from three systems differing only in their temperature. (The flow lines are schematic but display the essential features revealed in detailed calculations.)

Consider first the red dots which start from the nearest neighbour Ising model at its critical point. The ratio of correlation length to coarse-graining scale is reduced by a factor b at each step, but, since it starts infinite, it remains infinite after any finite number of steps. In this case we can in principle generate an unbounded number of dots, $\mathbf{K^{(1)}}, \mathbf{K^{(2)}}, \ldots, \mathbf{K^{(n)}}$, all of which lie in the critical surface. The simplest behaviour of such a sequence as n increases is to tend to a limit, K^*, say. In such a case

$$K_a^* = f_a(K^*) \qquad a = 1, 2..... \tag{60}$$

This point $\mathbf{K}^* \equiv K_1^*, K_2^*, \ldots$ is therefore a *fixed point* which lies in the critical surface.

By contrast, consider the same magnet as before, now at temperature T just greater than T_c, its couplings K_a, will be close to the first red dot (in fact they will be slightly smaller) and so will the effective couplings $K_a^{(1)}, K_2^{(2)}, \ldots$ of the corresponding coarse-grained systems. The new flow will therefore appear initially to follow the red dots towards the same fixed point. However, the flow must eventually move away from the fixed point because each coarse-graining now produces a model further from criticality. The resulting flow is represented schematically by one set of black dots. The other set of black dots shows the expected flow starting from the same magnet slightly below its critical temperature.

We are now in a position to understand both universality and scaling within this framework. We will suppose that there exists a single fixed point in the critical surface which sucks in all flows starting from a point in that surface. Then any system at its critical point will exhibit large-length scale physics (large-block spin behaviour) described by the single set of fixed point coupling constants. The uniqueness of this limiting set of coupling constants is the essence of critical point universality. It is, of course, the algebraic counterpart of the unique limiting spectrum of coarse-grained configurations, discussed in sec 7.5. Similarly the scale-invariance of the critical point configuration spectrum (viewed on large enough length scales) is expressed in the invariance of the couplings under iteration of the transformation (after a number of iterations large enough to secure convergence to the fixed point).

To understand the behaviour of systems near but not precisely at critically we must make a further assumption (again widely justified by explicit studies). The flow line stemming from any such system will, we have argued, be borne towards the fixed point before ultimately deviating from it after a number of iterations large enough to expose the system's noncritical character. We assume that (as indicated schematically in the streams of red and blue lines in fig. 16) the deviations lie along a single line through the fixed point, the direction followed along this line differing according to the sign of the temperature deviation $T - T_c$. Since any two sets of coupling constants on the line (on the same side of the fixed point) are related by a suitable coarse-graining operation, this picture implies that the large-length-scale physics of all near-critical systems differs only in the matter of a length scale. This is the essence of near-critical point universality.