

Table of contents

Welcome to the Complex Disordered Matter Course!



Overview

This course introduces you to the theoretical, computational and experimental aspects of the physics of complex disordered matter.

Complex disordered matter is the study of wide range of systems like **polymers**, **colloids**, **glasses**, **gels**, and **emulsions**, which lack long-range order but exhibit intricate behaviour. Colloids, suspensions of microscopic particles in a fluid, are useful for studying disordered structures due to their observable dynamics. Similarly, polymer systems can form amorphous solids or glasses when densely packed or cooled, showing solid-like rigidity despite their disordered structure. These materials often undergo phase transitions, such as demixing and crystallisation, and near these transitions, they can display critical phenomena with extensive fluctuations and correlations.

These various are examples of **soft matter**. In soft matter systems, the interplay between disorder, softness, and phase behavior leads to rich physical phenomena, particularly near critical points where even small changes in external conditions can trigger large-scale reorganizations and universal behaviour. Glasses, for instance, exhibit slow relaxation and memory effects, while colloidal systems may crystallize, phase separate, or become jammed depending on particle interactions and concentration. Understanding such behaviors involves studying how microscopic interactions and thermal fluctuations influence macroscopic properties, especially in non-equilibrium conditions. Through techniques like scattering, microscopy, rheology, and simulation, one can explore how disordered soft materials respond to stress, age, or undergo transitions—insights that are vital for applications in materials design, biotechnology, and beyond.

This course is organized into three interconnected parts, each offering a distinct perspective on the study of complex disordered matter.

- **Part 1: Unifying concepts** (Nigel Wilding) introduces the theoretical framework for rationalising complex disordered matter which is grounded in statistical mechanics and thermodynamics. We emphasize the theory of phase transitions, thermal fluctuations, critical phenomena, and stochastic dynamics—providing the essential theoretical tools

needed to describe and predict the behavior of soft and disordered systems.

- **Part 2: Complex disordered matter** (Francesco Turci) explores the phenomenology of key examples of complex disordered soft matter systems, including colloids, polymers, liquid crystals, glasses, gels, and active matter. These systems will be analyzed using the theoretical concepts introduced in Part 1, highlighting how disorder, interactions, and fluctuations shape their macroscopic behavior.
- **Part 3: Experimental techniques** (Adrian Barnes) focuses on the methods of microscopy, and scattering via x-rays, neutrons and light that are used to study complex disordered matter, offering insight into how their properties are measured and understood in real-world contexts.

In addition to theory and experiment, computer simulation plays a central role in soft matter research. This course includes a substantial coursework component consisting of two computational projects. These exercises will allow you to apply state-of-the-art simulation techniques to investigate the complex behavior of disordered systems, bridging theory and observation through hands-on exploration.

Delivery and format

- Detailed e-notes (accessible via Blackboard) can be viewed on a variety of devices. Pdf is also available.
- We will give ‘traditional’ lectures (Tuesdays, Wednesdays, Fridays) in which we use slides to summarise and explain the lecture content. Questions are welcome (within reason...)
- Try to read ahead in the notes, then come to lectures, listen to the explanations and then reread the notes.
- Rewriting the notes or slides to express your own thoughts and understanding, or annotating a pdf copy can help wire the material into your own way of thinking.
- There are problem classes (Thursdays) where you can try problem sheets and seek help. Lecturers may go over some problems with the class.
- The navigation bar on the left will allow you to access the lecture notes and problem sets.

Intended learning outcomes

The course will

- Introduce you to the qualitative features of a range of complex and disordered systems and the experimental techniques used to study them.
- Introduce you to a range of model systems and theoretical techniques used to elucidate the physics of complex disordered matter.
- Provide you with elementary computational tools to model complex disordered systems numerically and predict their properties.
- Allow you to apply your physics background to understand a variety of systems of interdisciplinary relevance.
- Connect with the most recent advances in the research on complex disordered matter.

Contact details

The course will be taught by

- Prof Nigel B. Wilding (unit director): nigel.wilding@bristol.ac.uk
- Dr Francesco Turci: F.Turci@bristol.ac.uk
- Dr Adrian Barnes: a.c.barnes@bristol.ac.uk

Questions and comments

If you have any questions about the course, please don't hesitate to contact the relevant lecturer, either by email (see above) or in a problems class.

Finally, this is a new course for 2025/26. If you find any errors or mistakes or something which isn't clear, please let us know by email, or fill in this anonymous form:

[Submit an error/mistake/query](#)

Recommended texts and literature

Needs to be tidied and unified.

Experimental textbooks?

One motivation for supplying you with detailed notes for this course course is the absence of a wholly ideal text book. However, it should be stressed that while these notes approach (in places) the detail of a book, 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 textbooks are:

A good book at the right level for the phase transitions and critical phenomena part of the course is

- **J.M. Yeomans: Statistical Mechanics of Phase Transitions**

A good book covering all aspects of this part of the course including non-equilibrium systems is

- **D. Chandler: Introduction to Modern Statistical Mechanics**

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

- **N Goldenfeld: Lectures on Phase Transitions and the Renormalization Group**
- **J.J. Binney, N.J. Dowrick, A.J.Fisher and M.E.J. Newman: The Theory of Critical Phenomena**

For revision on thermodynamics and statistical mechanics

- **F. Mandl: Statistical Physics.**

For Stochastic dynamics

- **N.G. van Kampen: Stochastic processes in Physics and Chemistry**

The best overall text for part 2 of the course is: R.A.L Jones, Soft Condensed Matter, Oxford University Press.

Additionally, the following more specialised texts should also be useful. They can be found in the University Library under the stated shelfmark.

Colloids

- (1) D.F.Evans, H.Wennerström: The Colloidal Domain - Where Physics, Chemistry, Biology, and Technology Meet. VCH Publishers (1994).
- (2) R.J.Hunter: Introduction to Modern Colloid Science. Oxford University Press (1993).
- (3) W.B.Russel, D.A.Saville, W.R.Schowalter: Colloidal Dispersions Cambridge University Press (1989).
- (4) D.H.Everett: Basic Principles of Colloid Science.

Royal Society of Chemistry Paperbacks (1988)

Polymers and surfactants

- (1) R.J. Young and P.A. Lovell: Introduction to polymers .
- (2) M. Doi: Introduction to polymer physics
- (3) J.Israelachvili, Intermolecular and Surface Forces, Academic Press (1992), Chs. 16 and 17

Glasses

- (1) J. Zarzycki; Glasses and the vitreous state. Cambridge University Press (1991).

Part I

Unifying concepts

Chapter 1

Introduction to phase behaviour and enhanced fluctuations

A phase transition can be defined as a macroscopic rearrangement of the internal constituents of a system in response to a change in the thermodynamic conditions to which they are subject. A wide variety of physical systems undergo such transitions. Understanding the properties of phase transitions is fundamental to the study of soft and complex matter, as these systems often exhibit rich and subtle transformations between different states of organization. Whether in colloidal suspensions, polymer blends, liquid crystals, or biological materials, phase transitions underpin a wide range of physical behaviours, from self-assembly and pattern formation to critical phenomena and dynamical arrest. By analysing how macroscopic phases emerge from microscopic interactions and external conditions, one gains crucial insight into the principles that govern structure, stability, and functionality in these intricate systems. As such, an understanding of phase transitions not only enriches theoretical understanding but also informs practical applications across materials science, biophysics, and nanotechnology. For these reasons we will devote a large proportion of this course to the study of phase transitions.

Two classic examples of systems displaying phase transitions are the ferromagnet and fluid systems. For the magnet, a key observable is the magnetisation defined as the magnetic moment per spin, given by $m = M/N$, with N the number of spins. m can be positive or negative, dependent on whether the spins are aligned ‘up’ or ‘down’. As the temperature of a ferromagnet is increased, its net magnetisation $|m|$ is observed to decrease smoothly, until at a certain temperature known as the critical temperature, T_c , it vanishes altogether (see left part of Figure ??). We define the magnetisation to be the *order parameter* of this phase transition.

One can also envisage applying a magnetic field H to the system which, depending on its sign (i.e. whether it is aligned (positive) or anti-aligned (negative) relative to the magnetisation axis), favours up or down spin states respectively, as shown schematically in Figure ?? (right part). Changing the sign of the magnetic field H for $T < T_c$ leads to a phase transition characterised by a discontinuous jump in m . We shall explore this behaviour in more detail in section 6.

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 number 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 and the two phases coalesce. These conditions of pressure and temperature serve to locate the critical point for the fluid. We define the density difference $\rho_{\text{liq}} - \rho_{\text{vap}}$ to be the order parameter for the liquid-gas phase transition. We shall meet order parameters for other,

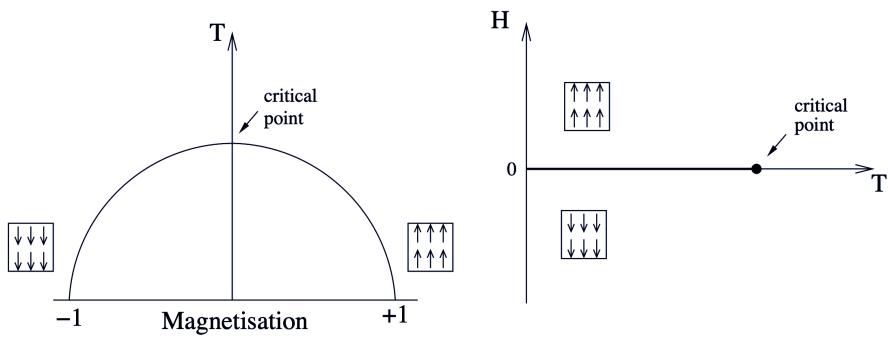


Figure 1.1: Phase diagram of a simple magnet (schematic). Left: magnetisation as a function of temperature for zero applied magnetic field, $H = 0$. Right: Applying a magnetic field that is aligned or antialigned with the direction of the magnetisation leads to a phase transition. The $H = 0$ axis at $T < T_c$ is the coexistence curve for which positive and negative magnetisations are equally likely.

more complex, systems in Section ??,

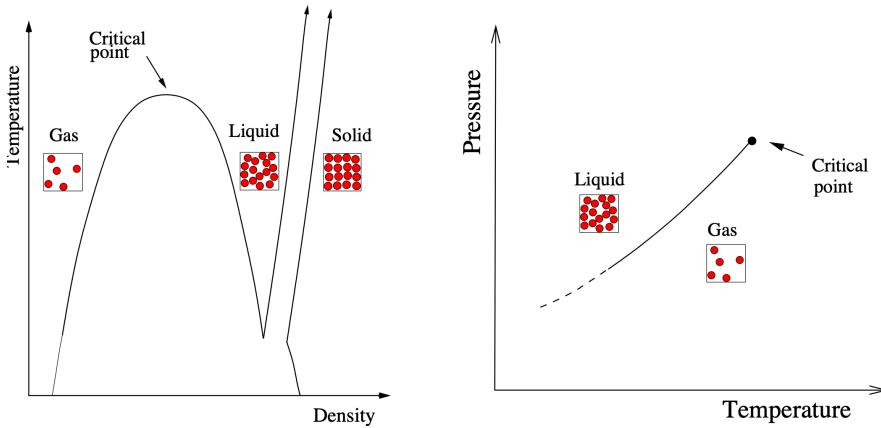


Figure 1.2: Phase diagram of a simple fluid (schematic)

In the vicinity of a critical point, a system displays a host of remarkable behaviors known as *critical phenomena*. Chief among these is the divergence of thermal response functions—such as specific heat, compressibility, or magnetic susceptibility—which signal an enhanced sensitivity to external perturbations. These singularities arise from the emergence of large-scale cooperative interactions among the system's microscopic constituents, as measured by a diverging *correlation length* (see Chapter ??). One visually striking manifestation of this is *critical opalescence*, particularly observed in fluids like CO_2 . As carbon dioxide nears its critical temperature and pressure, the distinction between its liquid and gas phases vanishes, giving rise to huge fluctuations in density. These fluctuations scatter visible light, rendering the fluid milky or opalescent. This scattering effect directly reflects the long-range correlations developing within the fluid. The movie below illustrates the effect as the critical temperature of CO_2 is approached from above. Note the appearance of a liquid-vapour interface (meniscus) as the system enters the two-phase region.

[Movies/critical_point_1.mp4](#)

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. Systems such as liquid mixtures, superconductors, liquid crystals, ferromagnets, antiferromagnets and molecular crystals may display identical behaviour near criticality. 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 in the 1970s, has enjoyed considerable success.

In the next few lectures, principal aspects of the contemporary theoretical viewpoint of phase transitions and critical phenomena will be reviewed. Mean field theories of phase transitions will be discussed and their inadequacies in the critical region will be exposed. The phenomenology of the critical region will be described including power laws, critical exponents and their relationship to scaling phenomena. These will be 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. Thereafter we will consider some foundational concepts in the dynamics of complex disordered matters. We shall look at the processes by which one phase transform into another and introduce differential equations that allow us to deal with the inherent stochasticity of thermal systems. The wider applicability of these unifying concepts to complex disordered systems such as colloids, polymers, liquid crystals and glasses will be covered in part 2 of the course.

Chapter 2

Key concepts for phase transitions

2.1 Observables and expectation values

In seeking to describe phase transition and critical phenomena, it is useful to have a quantitative measure of the difference between the phases: 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 suggests, 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. Referring back to [?@sec-canonical](#), the probability p_a that a physical system at temperature T will have a particular microscopic arrangement (alternatively referred to as a ‘configuration’ or ‘state’), labelled a , of energy E_a is

$$p_a = \frac{1}{Z} e^{-E_a/k_B T} \quad (2.1)$$

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

$$Z = \sum_a e^{-E_a/k_B T} \quad (2.2)$$

where the sum extends over all possible microscopic arrangements.

These equations assume that physical system evolves rapidly (on the timescale of typical observations) amongst all its allowed arrangements, sampling them with the probabilities Equation ?? 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:

$$\bar{O} = \frac{1}{Z} \sum_a O_a e^{-E_a/k_B T} \quad (2.3)$$

Sums like Equation ?? 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.

$$Q = \frac{1}{Z} \sum_a Q_a e^{-E_a/k_B T} \quad (2.4)$$

It is clear from Equation ?? 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 Equation ?? 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 Equation ?? 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. Specifically, the bridge between the two disciplines is provided by the following equation

$$F = -k_B T \ln Z \quad (2.5)$$

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, utilizing Equation ??, one can readily verify (try it as an exercise!) that the average internal energy is given by

$$\overline{E} = -\frac{\partial \ln Z}{\partial \beta},$$

where $\beta = (k_B T)^{-1}$.

The relationship between other thermodynamic quantities and derivatives of the free energy are given in fig. Figure ??

2.2 Correlations

2.2.1 Spatial correlations

The two-point connected correlation function measures how fluctuations at two spatial points are statistically related. For a scalar field $\phi(\vec{R})$, which could represent eg. the local magnetisation m in a magnet at position vector \vec{R} , or the local particle number density ρ in a fluid, it is defined as:

$$C(r) = \langle \phi(\vec{R})\phi(\vec{R} + \vec{r}) \rangle - \langle \phi(\vec{R}) \rangle^2,$$

$$\begin{array}{c}
Z = \sum_a e^{-E_a/k_B T} \\
\downarrow \\
F = -k_B T \ln Z \\
\swarrow \quad \searrow \\
\bar{E} = - \left(\frac{\partial(\beta F)}{\partial \beta} \right)_H \quad \bar{M} = - \left(\frac{\partial F}{\partial H} \right)_T \quad (M = mN) \\
\downarrow \quad \downarrow \\
C_H = \left(\frac{\partial \bar{E}}{\partial T} \right)_H \quad \chi_T = \left(\frac{\partial m}{\partial H} \right)_T
\end{array}$$

Figure 2.1: Relationships between the partition function and thermodynamic observables

where $\langle \cdot \rangle$ denotes an ensemble or spatial average over all \vec{R} , and $r = |\vec{r}|$ is the spatial separation between the two points.

$C(r)$ quantifies the spatial extent over which field values are correlated and in homogeneous and isotropic systems, it depends only on the separation r .

If $C(r)$ decays quickly, we say that correlations are short-ranged. Typically this occurs well away from criticality and takes the form of exponential decay

$$C(r) \sim e^{-r/\xi}$$

where the correlation length ξ is the characteristic scale over which correlations decay.

Near a critical point $C(r)$ decays more slowly - in a power-law fashion - and correlations are long-ranged.

$$C(r) \sim r^{-(d-2+\eta)}$$

where d is the spatial dimension and η is a critical exponent.

In isotropic fluids and particle systems, a closely related and more directly measurable quantity (particularly in simulations) is the **radial distribution function** $g(r)$, which describes how particle density varies as a function of distance from a reference particle. For such systems, the two-point correlation function of the number density field $\rho(\vec{r})$ is related to $g(r)$ as follows:

$$g(r) = 1 + \frac{C(r)}{\rho^2},$$

where ρ is the average number density. This relation shows that $g(r)$ encodes the same spatial correlations as $C(r)$, but in a form that is more natural for discrete particle systems. Note that

by definition $g(r) \rightarrow 1$ in the absence of correlations ie. when $C(r) = 0$. This is typically the case for $r \gg \xi$.

Experimentally one doesn't typically have direct access to $C(r)$, but rather its Fourier transform known as the **structure factor**

$$S(k) = \int d^d r e^{-i\vec{k} \cdot \vec{r}} C(r),$$

where k is the scattering wavevector.

In equilibrium:

- For short-range correlations (finite ξ), $S(k)$ typically has a Lorentzian form:

$$S(k) \sim \frac{1}{k^2 + \xi^{-2}}.$$

- At criticality (where $\xi \rightarrow \infty$), $S(k)$ follows a power law:

$$S(k) \sim k^{-2+\eta}.$$

This relation enables the extraction of ξ from experimental or simulation data, especially via scattering techniques.

2.2.2 Temporal correlations

Consider a thermodynamic variable x with zero mean that fluctuates over time. Examples include the local magnetization in a magnetic system or the local density in a fluid. Here, x represents a deviation from the average value — a fluctuation.

We're interested in how such fluctuations are correlated over time when the system is in thermal equilibrium. For instance, if x is positive at some time t , it's more likely to remain positive shortly after.

These temporal correlations are characterized by the two-time correlation function (also known as an auto-correlation function):

$$\langle x(\tau)x(\tau+t) \rangle$$

In equilibrium, the correlation function must be independent of the starting time τ . Therefore, we define:

$$\langle x(\tau)x(\tau+t) \rangle = M_{xx}(t)$$

That is, $M_{xx}(t)$ depends only on the time difference t .

We typically expect $M_{xx}(t)$ to decay exponentially over a characteristic correlation time t_c :

$$M_{xx}(t) \sim \exp(-t/t_c)$$

This exponential decay reflects how the memory of fluctuations fades with time.

Now consider two different fluctuating variables, x and y (e.g., local magnetizations at different positions). Their cross-correlation function is defined as:

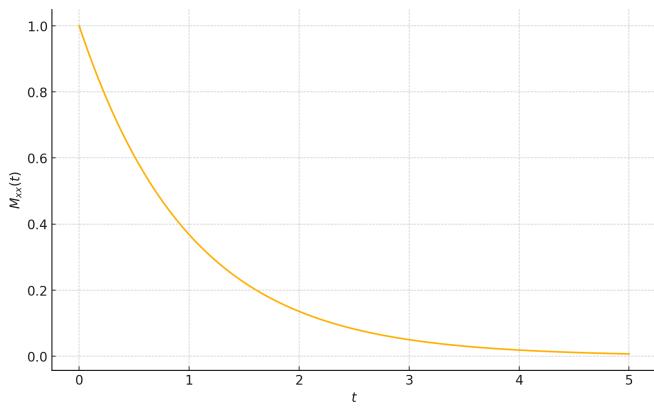


Figure 2.2: Sketch of $M_{xx}(t)$ against t

$$\langle x(\tau)y(\tau+t) \rangle = M_{xy}(t)$$

This defines the elements of a dynamic correlation matrix, of which $M_{xx}(t)$ is the diagonal.

Chapter 3

The approach to criticality

It is a matter of experimental fact that the approach to criticality in a given system is characterized 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 a *power* of the reduced (dimensionless) temperature $t \equiv (T - T_c)/T_c$:

$$\chi \equiv \frac{\partial M}{\partial H} \propto t^{-\gamma} \quad (H = 0) \quad (3.1)$$

(where $M = mN$),

$$C_H \equiv \frac{\partial E}{\partial T} \propto t^{-\alpha} \quad (H = \text{constant}) \quad (3.2)$$

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 .

$$\xi \propto t^{-\nu} \quad (T > T_c, H = 0) \quad (3.3)$$

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 approached as a function of temperature:

$$m \propto t^\beta \quad (T < T_c, H = 0) \quad (3.4)$$

(here the symbol β , is not to be confused with $\beta = 1/k_B T$ – this unfortunately is the standard notation.)

Finally, as a function of magnetic field:

$$m \propto h^{1/\delta} \quad (T = T_c, H > 0). \quad (3.5)$$

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

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

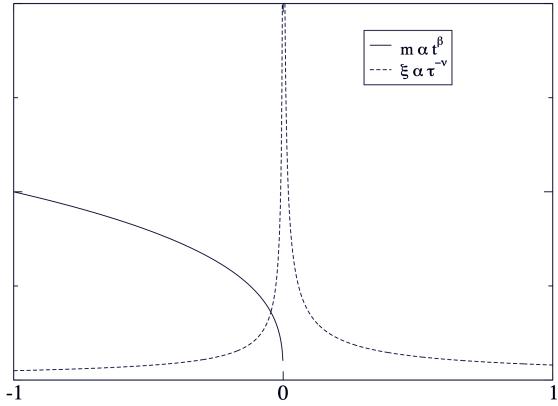


Figure 3.1: Singular behaviour of the correlation length and order parameter in the vicinity of the critical point as a function of the reduced temperature t .

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.

More about the scientists mentioned in this chapter:

[Kenneth Wilson](#)

Chapter 4

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 realizations 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.

4.1 The 2D Ising model

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

$$H_I = -J \sum_{\langle ij \rangle} s_i s_j - H \sum_i s_i \quad (4.1)$$

where $J > 0$ measures the strength of the coupling between spins and the sum extends over nearest neighbour spins s_i and s_j , i.e it is a sum of the bonds of the lattice. 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:

$$m = \frac{1}{N} \left\langle \sum_i s_i \right\rangle ,$$

where $\langle \cdot \rangle$ means an average over typical configurations corresponding to the prescribed value of $J/k_B T$.

The fact that the Ising model displays a phase transition was argued in Chapter ???. Thus at low temperatures for which there is little thermal disorder, there is a preponderance of aligned spins and hence a net spontaneous magnetic moment (ie. the system is ferromagnetic). 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 (ie. the system is paramagnetic). These trends are visible in configurational snapshots from computer simulations of the 2D Ising model (see Figure ??). Although each spin interacts only with its nearest neighbours, the phase transition occurs due to cooperative effects among a large number of spins.

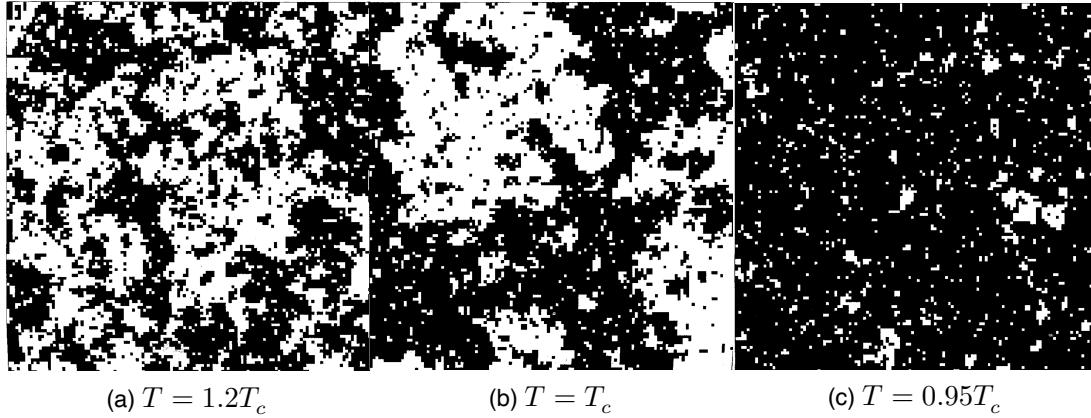


Figure 4.1: Configurations of the 2d Ising model. The patterns depict typical arrangements of the spins (white=+1, black=-1) generated in a computer simulation of the Ising model on a square lattice of $N = 512$ sites, at temperatures (from left to right) of $T = 1.2T_c$, $T = T_c$, and $T = 0.95T_c$. In each case only a portion of the system containing 128 sites is 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 Monte Carlo simulation of the Ising model](#) demonstrates the phenomenology. By altering the temperature you will be able to observe for yourself how the typical 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. We shall explore these ideas in more detail later in the course.

4.2 Exact solutions: the one dimensional Ising chain

One might well ask why the 2D Ising model is the simplest model to exhibit a phase transition. What about the one-dimensional Ising model (ie. spins on a line)? In fact in one dimension, the Ising model can be solved exactly. It turns out that the system is paramagnetic for all $T > 0$, so there is no phase transition at any finite temperature. To see this, consider the ground state of the system in zero external field. This will have all spins aligned the same way (say up), and hence be ferromagnetic. Now consider a configuration with a various “domain walls” dividing spin up and spin down regions:

Instead of considering the underlying spin configurations, we shall describe the system in terms of the statistics of its domain walls. The energy cost of a wall is $\Delta = 2J$, independent of position. Domain walls can occupy the bonds of the lattice, of which there are $N - 1$. Moreover, the walls are noninteracting, except that you cannot have two of them on the same bond. (Check through these ideas if you are unsure.)

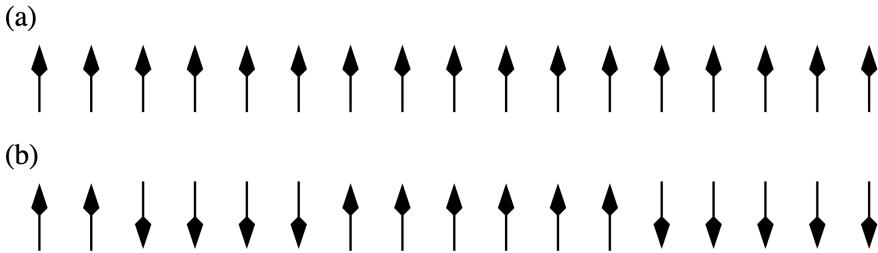


Figure 4.2: (a) Schematic of an Ising chain at $T = 0$. (b) At a small finite temperature the chain is split into domains of spins ordered in the same direction. Domains are separated by notional domain “walls”, which cost energy $\Delta = 2J$. Periodic boundary conditions are assumed.

In this representation, the partition function involves a count over all possible domain wall arrangements. Since the domain walls are non interacting (eg it doesn't cost energy for one to move along the chain) we can calculate Z by considering the partition function associated with a single domain wall being present or absent on some given bond, and then simply raise to the power of the number of bonds:

$$Z = Z_1^{N-1}$$

where

$$Z_1 = e^{\beta J} + e^{\beta(J-\Delta)} = e^{\beta J}(1 + e^{-\beta\Delta})$$

is the domain wall partition function for a single bond and represent the sum over the two possible states: domain wall absent or present. Then the free energy per bond of the system is

$$\beta f \equiv \beta F/(N - 1) = -\ln Z_1 = -\beta J - \ln(1 + e^{-\beta\Delta})$$

The first term on the RHS is simply the energy per spin of the ferromagnetic (ordered) phase, while the second term arises from the free energy of domain walls. Clearly for any finite temperature (ie. for $\beta < \infty$), this second term is finite and negative. Hence the free energy will always be lowered by having a finite concentration of domain walls in the system. Since these domain walls disorder the system, leading to a zero average magnetisation, the 1D system is paramagnetic for all finite temperatures. *Exercise:* Explain why this argument works only in 1D.

The animation below lets you see qualitatively how the typical number of domain walls varies with temperature. If you'd like to explore more quantitatively, a python code performing a Monte Carlo simulation is available. You will learn about Monte Carlo simulation in the coursework and in later parts of the course.

```
#Monte Carlo simulation of the 1d Ising chain with periodic boundary conditions
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.animation import FuncAnimation
from matplotlib.widgets import Slider

# Number of spins
N = 20

# Initialize spins (+1 or -1)
spins = np.random.choice([-1, 1], size=N)
```

```

# Initial temperature
T = 2.0

# Set up figure and axis for the spins
fig, ax = plt.subplots(figsize=(10, 2))
plt.subplots_adjust(bottom=0.25) # make room for slider
ax.set_xlim(-0.5, N - 0.5)
ax.set_ylim(-1, 1)
ax.axis('off')

# Create text objects for each spin
texts = []
for i in range(N):
    arrow = '↑' if spins[i] == 1 else '↓'
    t = ax.text(i, 0, arrow, fontsize=24, ha='center', va='center')
    texts.append(t)

def update(frame):
    """Perform Metropolis updates over all spins, then refresh display."""
    global spins, T
    for _ in range(N):
        i = np.random.randint(N)
        left = spins[(i - 1) % N]
        right = spins[(i + 1) % N]
        deltaE = 2 * spins[i] * (left + right)
        # Metropolis criterion ensures configurations appear with the correct Boltzmann probability
        if deltaE < 0 or np.random.rand() < np.exp(-deltaE / T):
            spins[i] *= -1
    # Update arrows on screen
    for idx, t in enumerate(texts):
        t.set_text('↑' if spins[idx] == 1 else '↓')
    return texts

# Create the animation with caching disabled and blit turned off
ani = FuncAnimation(
    fig,
    update,
    interval=200,
    blit=False,
    cache_frame_data=False
)

# Add a temperature slider
ax_T = plt.axes([0.2, 0.1, 0.6, 0.03], facecolor='lightgray')
slider_T = Slider(ax_T, 'Temperature T', 0.1, 5.0, valinit=T)

def on_T_change(val):
    """Callback to update T when the slider changes."""
    global T
    T = val

slider_T.on_changed(on_T_change)

```

```
# Show the plot (ani is kept in scope so it won't be deleted)
plt.show()
```

Temperature T =

2.0

4.2.1 More general 1D spins systems: transfer matrix method

Generally speaking one-dimensional systems lend themselves to a degree of analytic tractability not found in most higher dimensional models. Indeed for the case of a 1-d assembly of N spins each having m discrete energy states, and in the presence of a magnetic field, it is possible to reduce the evaluation of the partition function to the calculation of the eigenvalues of a matrix—the so called transfer matrix.

Let us start by assuming that the assembly has cyclic boundary conditions, then the total energy of configuration $\{s\}$ is

$$\begin{aligned} H(\{s\}) &= - \sum_{i=1}^N (J s_i s_{i+1} + H s_i) \\ &= - \sum_{i=1}^N (J s_i s_{i+1} + H(s_i + s_{i+1})/2) \\ &= \sum_{i=1}^N E(s_i, s_{i+1}) \end{aligned}$$

where we have defined $E(s_i, s_{i+1}) = -J s_i s_{i+1} - H(s_i + s_{i+1})/2$.

Now the partition function may be written

$$\begin{aligned} Z_N &= \sum_{\{s\}} \exp(-\beta H(\{s\})) \\ &= \sum_{\{s\}} \exp(-\beta [E(s_1, s_2) + E(s_2, s_3) + \dots + E(s_N, s_1)]) \\ &= \sum_{\{s\}} \exp(-\beta E(s_1, s_2)) \exp(-\beta E(s_2, s_3)) \dots \exp(-\beta E(s_N, s_1)) \\ &= \sum_{i,j,\dots,l=1}^m V_{ij} V_{jk} \dots V_{li} \end{aligned}$$

where the $V_{ij} = \exp(-\beta E_{ij})$ are elements of an $m \times m$ matrix \mathbf{V} , known as the transfer matrix (i, j, k etc are dummy indices that run over the matrix elements). You should see that the sum over the product of matrix elements picks up all the terms in the partition function and therefore Equation ?? is an alternative way of writing the partition function.

The reason it is useful to transform to a matrix representation is that it transpires that the sum over the product of matrix elements in Equation ?? is simply just the trace of \mathbf{V}^N (check this yourself for a short periodic chain), given by the sum of its eigenvalues:-

$$Z_N = \lambda_1^N + \lambda_2^N + \dots \lambda_m^N$$

For very large N , this expression simplifies further because the largest eigenvalue λ_1 dominates the behaviour since $(\lambda_2/\lambda_1)^N$ vanishes as $N \rightarrow \infty$. Consequently in the thermodynamic limit one may put $Z_N = \lambda_1^N$ and the problem reduces to identifying the largest eigenvalue of the transfer matrix.

Specializing to the case of the simple Ising model in the presence of an applied field H , the transfer matrix takes the form

$$\mathbf{V}(H) = \begin{pmatrix} e^{\beta(J+H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-H)} \end{pmatrix}$$

This matrix has two eigenvalues which can be readily calculated in the usual fashion as the roots of the characteristic polynomial $|\mathbf{V} - \lambda \mathbf{I}|$. They are

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta H) \pm \sqrt{e^{2\beta J} \sinh^2 \beta H + e^{-2\beta J}}.$$

Hence the free energy per spin $f = -k_B T \ln \lambda_+$ is

$$f = -k_B T \ln \left[e^{\beta J} \cosh(\beta H) + \sqrt{e^{2\beta J} \sinh^2 \beta H + e^{-2\beta J}} \right].$$

The Ising model in 2D can also be solved exactly, as was done by Lars Onsager in 1940. The solution is extremely complicated and is regarded as one of the pinnacles of statistical mechanics. In 3D no exact solution is known.

Chapter 5

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 phase coexistence parameters, critical exponents and the critical 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.

5.1 Mean field solution of the Ising model

Let us look for a mean field expression for the free energy of the Ising model whose Hamiltonian is given in Equation ?? . Write

$$s_i = \langle s_i \rangle + (s_i - \langle s_i \rangle) = m + (s_i - m) = m + \delta s_i$$

Then

$$\begin{aligned} H_I &= -J \sum_{\langle i,j \rangle} [m + (s_i - m)][m + (s_j - m)] - H \sum_i s_i \\ &= -J \sum_{\langle i,j \rangle} [m^2 + m(s_i - m) + m(s_j - m) + \delta s_i \delta s_j] - H \sum_i s_i \\ &= -J \sum_i (qms_i - qm^2/2) - H \sum_i s_i - J \sum_{\langle i,j \rangle} \delta s_i \delta s_j \end{aligned}$$

where in the last line we have used the fact that when for each site i we perform the sum $\sum_{\langle i,j \rangle}$ over bonds of a quantity which is independent of s_j , then the result is just the number of bonds per site times that quantity. Since the number of bonds on a lattice of N sites of coordination q

is $Nq/2$ (because each bond is shared between two sites), there are therefore $q/2$ bonds per site.

Now the mean field approximation is to ignore the last term in the last line of Equation ?? giving the configurational energy as

$$H_{mf} = - \sum_i H_{mf} s_i + NqJm^2/2$$

with $H_{mf} \equiv qJm + H$ the “mean field” seen by spin s_i . As all the spins are decoupled (independent) in this approximation we can write down the partition function, which follows by taking the partition function for a single spin (by summing the Boltzmann factor for $s_i = \pm 1$) and raising to the power N to find

$$Z = e^{-\beta qJm^2N/2} [2 \cosh(\beta(qJm + H))]^N$$

The free energy follows as

$$F(m) = NJqm^2/2 - Nk_B T \ln[2 \cosh(\beta(qJm + H))].$$

and the magnetisation as

$$m = -\frac{1}{N} \frac{\partial F}{\partial H} = \tanh(\beta(qJm + H))$$

To find $m(H, T)$, we must numerically solve this last equation self consistently.

Note that we can obtain m in a different way. Consider some arbitrary spin, s_i say. Then this spin has an energy $H_{mf}(s_i)$. Considering this energy for both cases $s_i = \pm 1$ and the probability $p(s_i) = e^{-\beta H_{mf}(s_i)}/Z$ of each, we have that

$$\langle s_i \rangle = \sum_{s_i=\pm 1} s_i p(s_i)$$

but for consistency, $\langle s_i \rangle = m$. Thus

$$\begin{aligned} m &= \sum_{s_i=\pm 1} s_i p(s_i) \\ &= \frac{e^{\beta(qJm+H)} - e^{-\beta(qJm+H)}}{e^{\beta(qJm+H)} + e^{-\beta(qJm+H)}} \\ &= \tanh(\beta(qJm + H)) \end{aligned}$$

as before.

5.2 Spontaneous symmetry breaking

This mean field analysis reveals what is happening in the Ising model near the critical temperature T_c . Figure ?? shows sketches for $\beta F(m)/N$ as a function of temperature, where for simplicity we restrict attention to $H = 0$. In this case $F(m)$ is symmetric in m . Moreover, at high T , the entropy dominates and there is a single minimum in $F(m)$ at $m = 0$. As T is

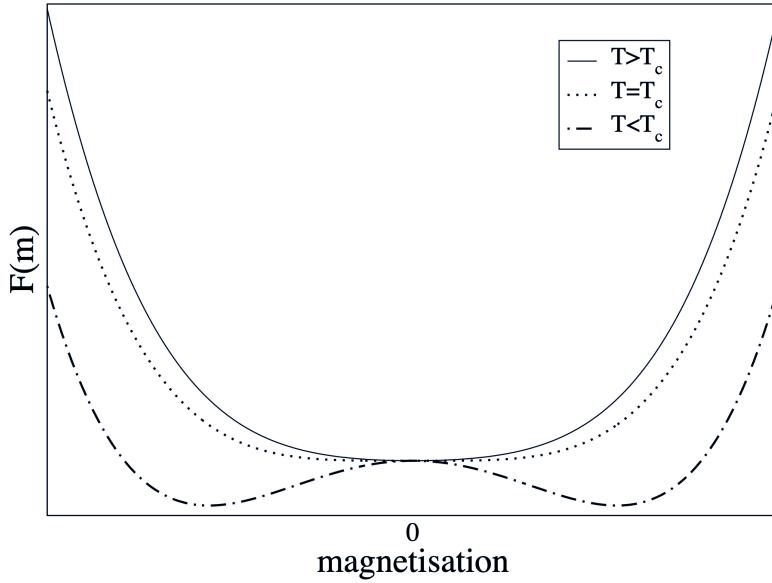


Figure 5.1: Schematic of the form of the free energy for a critical, subcritical and supercritical temperature

lowered, there comes a point ($T = T_c = qJ/k_B$) where the curvature of $F(m)$ at the origin changes sign; precisely at this point

$$\frac{\partial^2 F}{\partial m^2} = 0.$$

At lower temperature, there are instead two minima at nonzero $m = \pm m^*$, where the *equilibrium magnetisation* m^* is the positive root (calculated explicitly below) of

$$m^* = \tanh(\beta J q m^*) = \tanh\left(\frac{m^* T_c}{T}\right)$$

The point $m = 0$ which remains a root of this equation, is clearly an unstable point for $T < T_c$ (since F has a maximum there).

This is an example of spontaneous symmetry breaking. In the absence of an external field, the Hamiltonian (and therefore the free energy) is symmetric under $m \rightarrow -m$. Accordingly, one might expect the actual state of the system to also show this symmetry. This is true at high temperature, but spontaneously breaks down at low ones. Instead there are a pair of ferromagnetic states (spins mostly up, or spins mostly down) which – by symmetry– have the same free energy, lower than the unmagnetized state.

5.3 Phase diagram

The resulting zero-field magnetisation curve $m(T, H = 0)$ looks like Figure ??.

This shows the sudden change of behaviour at T_c (phase transition). For $T < T_c$ it is arbitrary which of the two roots $\pm m^*$ is chosen; typically it will be different in different parts of the sample (giving macroscopic “magnetic domains”). But this behaviour with temperature is *qualitatively modified* by the presence of a field H , however small. In that case, there is always a slight magnetization, even far above T_c and the curves becomes smoothed out, as shown. There is

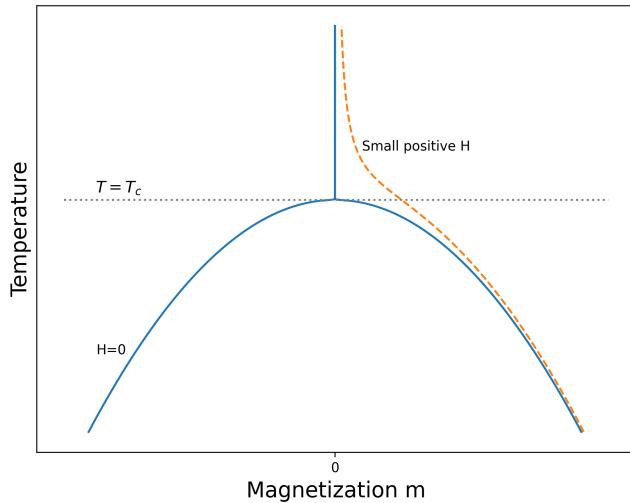


Figure 5.2: Phase diagram of a simple magnet in the m - T plane.

no doubt which root will be chosen, and no sudden change of the behaviour (no phase transition). Spontaneous symmetry breaking does not occur, because the symmetry is already broken by H . (The curve $F(m)$ is lopsided, rather than symmetrical about $m = 0$.)

On the other hand, if we sit below T_c in a positive field (say) and gradually reduce H through zero so that it becomes negative, there is a very sudden change of behaviour at $h = 0$: the equilibrium state jumps discontinuously from $m = m^*$ to $m = -m^*$.

This is called a first order phase transition as opposed to the “second order” or continuous transition that occurs at T_c in zero field. The definitions are:

First order transition: magnetisation (or similar order parameter) depends discontinuously on a field variable (such as h or T).

Continuous transition: Change of functional form, but no discontinuity in m ; typically, however, $(\partial m / \partial T)_h$ (or similar) is either discontinuous, or diverges with an integrable singularity.

In this terminology, we can say that the phase diagram of the magnet in the H , T plane shows a line of first order phase transitions, terminating at a continuous transition, which is the critical point.

🔥 Aside on Quantum Criticality

In some magnetic systems such as $CePd_2Si_2$, one can, by applying pressure or altering the chemical composition, depress the critical temperature all the way to absolute zero! This may seem counterintuitive, after all at $T = 0$ one should expect perfect ordering, not the large fluctuations that accompany criticality. It turns out that the source of the fluctuations that drive the system critical is zero point motion associated with the Heisenberg uncertainty principle. Quantum criticality is a matter of ongoing active research, and open questions concern the nature of the phase diagrams and the relationship to superconductivity. Although the subject goes beyond the scope of this course, there is an accessible article [here](#) if you want to learn more.

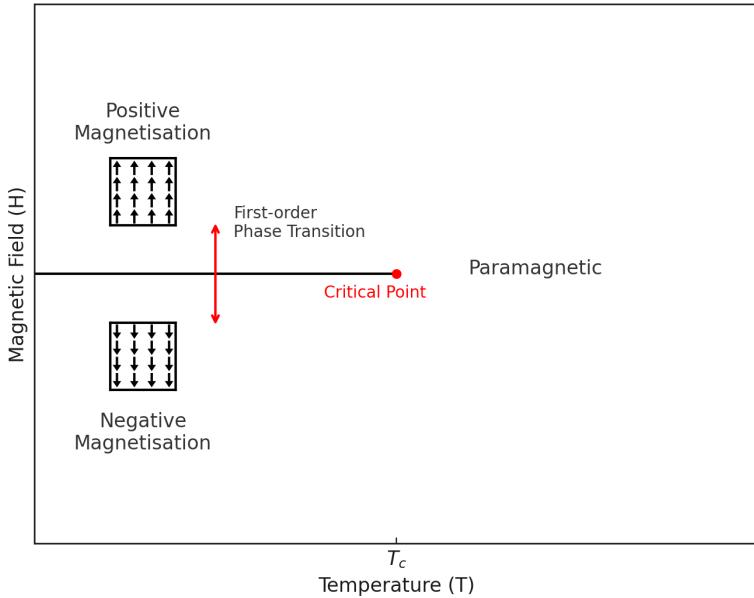


Figure 5.3: Phase diagram of a simple magnet in the H - T plane.

5.4 A closer look: critical exponents

Let us now see how we can calculate critical exponents within the mean field approximation.

5.4.1 Zero H solution and the order parameter exponent

In zero field

$$m = \tanh\left(\frac{mT_c}{T}\right)$$

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 $\beta = (k_B T)^{-1}$ yields

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

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

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

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

$$m^2 \simeq -3t$$

Hence

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

This result implies that (within the mean field approximation) the critical exponent $\beta = 1/2$.

5.4.2 Finite (but small) field solution: the susceptibility exponent

In a finite, but small field we can expand Equation ?? thus:

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

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

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

Hence near T_c

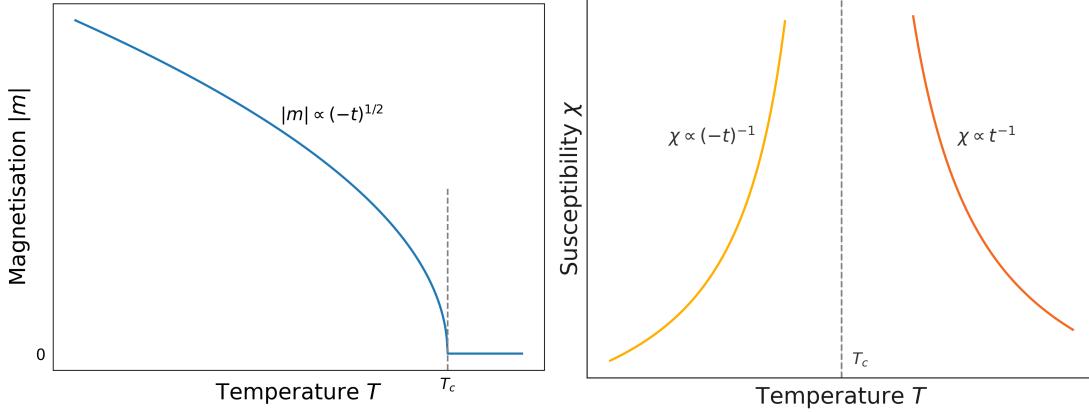
$$\chi = \frac{1}{k_B T_c} \left(\frac{1}{t + m^2} \right)$$

Then using the results of Equation ??

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

where one has to take the non-zero value for m below T_c to ensure +ve χ , i.e. thermodynamic stability. This result implies that (within the mean field approximation) the critical exponent $\gamma = 1$.

The schematic behaviour of the Ising order parameter and susceptibility are shown in Figure ?? (a) and (b) respectively.



(a) Mean field behaviour of the Ising magnetisation (schematic)
(b) Mean field behaviour of the Ising susceptibility (schematic)

5.5 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 power series expansion of the order parameter ϕ :

$$F(\phi) = \sum_{i=0}^{\infty} a_i \phi^i$$

The equilibrium value of ρ is that which minimises the Landau free energy.

🔥 A note on order parameters

We have already seen examples of these in earlier sections, e.g., for the liquid-gas transition this was

$\rho_{liq} - \rho_{gas}$: difference in density of two coexisting phases,

while for the Ising magnet it is the magnetisation m . Both quantities vanish at the critical point. These are examples of *scalar* order parameters – a single number is required to represent the degree of order ($n = 1$).

In the absence of a symmetry-breaking field, the Landau free-energy density f_L must have symmetry $f_L(-\phi) = f_L(\phi)$ (Ising case).

For some other systems, n component vectors are required in order to represent the order:

$$\phi = (\phi_1, \phi_2, \dots, \phi_n)$$

Then $f_L(\phi)$ should be symmetric under $O(n)$ rotations in n -component ϕ -space.

The table below lists examples of order parameters for various physical systems.

| Physical System | Order Parameter φ | Symmetry Group |
|------------------------------|-----------------------------|----------------|
| Uniaxial (Ising) ferromagnet | Magnetisation per spin, m | $O(1)$ |

| | | |
|--|--|--------|
| Fluid (liquid-gas) | Density difference, $\rho - \rho_c$ | $O(1)$ |
| Liquid mixtures | Concentration difference, $c - c_c$ | $O(1)$ |
| Binary (AB) alloy (e.g., β -brass) | Concentration of one of the species, c | $O(1)$ |
| Isotropic (vector) ferromagnet | n -component magnetisation, $\mathbf{m} = (m_1, m_2, \dots, m_n)$ | $O(n)$ |
| | $n = 2$: xy model | $O(2)$ |
| | $n = 3$: Heisenberg model | $O(3)$ |
| Superfluid He ⁴ | Macroscopic condensate wavefunction, Ψ | $O(2)$ |
| Superconductor (s-wave) | Macroscopic condensate wavefunction, Ψ | $O(2)$ |
| Nematic liquid crystal | Orientational order, $\langle P_2(\cos \theta) \rangle$ | |
| Smectic A liquid crystal | 1-dimensional periodic density | |
| Crystal | 3-dimensional periodic density | |

Notes:

- In **superfluid** 4He the order parameter is

$$\Psi = |\Psi| e^{i\theta},$$

the *complex wavefunction* of the macroscopic condensate. Both the amplitude $|\Psi|$ and phase θ must be specified, so this corresponds to $n = 2$.

Superconductors also correspond to $n = 2$.

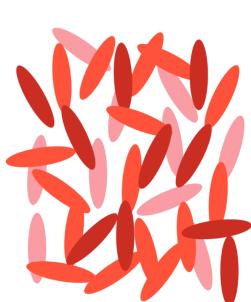
- In a **nematic** liquid crystal, the *orientational* order parameter is

$$\langle P_2(\cos \theta) \rangle \equiv \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle,$$

where θ is the angle a molecule makes with the average direction of the long axes of the molecules (known as the *director* \hat{n}). Rotational symmetry is broken. For the case of an n component vector, the free energy should be a function of:

$$\phi^2 \equiv |\phi|^2 = \phi_1^2 + \phi_2^2 + \dots + \phi_n^2 = \sum_{i=1}^n \phi_i^2$$

in the absence of a symmetry breaking field. Rotational symmetry is incorporated into the theory.



(a) Schematic of the isotropic liquid phase of a system of elongated molecules.



(b) Schematic of the nematic liquid phase of a system of elongated molecules. This phase has uniaxial ordering.

Figure 5.5: Isotropic and uniaxially ordered (nematic) phases of liquid crystal molecules.

To exemplify the approach, let us specialise to the case of a ferromagnet where $\phi = m$, the magnetisation and write the Landau free energy as

$$F(m) = F_0 + a_2 m^2 + a_4 m^4 \quad (5.1)$$

Here only the terms compatible with the order parameter symmetry are included in the expansion and we truncate the series at the 4th power because this is all that is necessary to yield the essential phenomenology. 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 $m \rightarrow -m$, and so do not feature.

One can understand how the Landau free energy can give rise to a critical point and coexistence values of the magnetisation, by plotting $F(m)$ for various values of a_2 with a_4 assumed positive (which ensures that the magnetisation remains bounded). This is shown in the following movie:

[Movies/landau_free_energy_evolution.mp4](#)

The situation is qualitatively similar to that discussed in Section ???. Thermodynamics tells us that the system adopts the state of lowest free energy. From the movie, 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 $m = \pm m^*$. $a_2 = 0$ corresponds to the critical point which marks the border between the ordered and disordered phases. Note that it is an inflexion point, so has $\frac{d^2 F}{dm^2} = 0$.

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

$$a_2 = \tilde{a}_2 t$$

where t is the reduced temperature. Thus we see that the trajectory of the minima as a function of $a_2 < 0$ in the above movie effective traces out the coexistence curve in the $m - T$ plane.

We can now attempt to calculate critical exponents. Restricting ourselves first to the magnetisation exponent β defined by $m = t^\beta$, we first find the equilibrium magnetisation, corresponding to the minimum of the Landau free energy:

$$\frac{dF}{dm} = 2\tilde{a}_2 tm + 4a_4 m^3 = 0 \quad (5.2)$$

which implies

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

so $\beta = 1/2$, which is again a mean field result.

Likewise we can calculate the effect of a small field H if we sit at the critical temperature T_c . Since $a_2 = 0$, we have

$$F(m) = F_0 + a_4 m^4 - Hm$$

$$\frac{\partial F}{\partial m} = 0 \Rightarrow m(H, T_c) = \left(\frac{H}{4a_4} \right)^{1/3}$$

or

$$H \sim m^\delta \quad \delta = 3$$

which defines a second critical exponent.

Note that at the critical point, a small applied field causes a very big increase in magnetisation; formally, $(\partial m / \partial H)_T$ is infinite at $T = T_c$.

A third critical exponent can be defined from the magnetic susceptibility at zero field

$$\chi = \left(\frac{\partial m}{\partial H} \right)_{T,V} \sim |T - T_c|^{-\gamma}$$

Exercise: Show that the Landau expansion predicts $\gamma = 1$.

Finally we define a fourth critical exponent via the variation of the heat capacity (per site or per unit volume) C_H , in fixed external field $H = 0$:

$$C_H \sim |T - T_c|^{-\alpha}$$

By convention, α is defined to be positive for systems where there is a *divergence* of the heat capacity at the critical point (very often the case). The heat capacity can be calculated from

$$C_H = -T \frac{\partial^2 F}{\partial T^2}$$

From the minimization over m @#eq-minimize one finds (*exercise:* check this)

$$\begin{aligned} F &= 0 \quad T > T_c \\ F &= -a_2^2/4a_4 \quad T < T_c \end{aligned}$$

Using the fact that a_2 varies linearly with T , we have

$$\begin{aligned} C_H &= 0 \quad T \rightarrow T_c^+ \\ C_H &= \frac{T\tilde{a}_2^2}{2a_4} \quad T \rightarrow T_c^- , \end{aligned}$$

which is actually a step discontinuity in specific heat. Since for positive α the heat capacity is divergent, and for negative α it is continuous, this behaviour formally corresponds to $\alpha = 0$

5.6 Shortcomings of mean field theory

While mean field theories provide a useful route to understanding qualitatively the phenomenology of phase transitions, in real ferromagnets, as well as in more sophisticated theories, the critical exponents 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. In later sections we shall start to take the first steps to including the effects of long range correlations.

Table 5.2: Comparison of true Ising critical exponents with their mean field theory predictions in a number of dimensions.

| | Mean Field | $d = 1$ | $d = 2$ | $d = 3$ |
|-----------------------------------|---------------|----------|---------------|-------------------|
| Critical temperature $k_B T/qJ$ | 1 | 0 | 0.5673 | 0.754 |
| 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 |

Chapter 6

The Lattice Gas model

A crude representation of a fluid is the lattice gas model. Here particles can occupy the sites of a hypercubic lattice. The occupancy of a site i is specified by the variable $c_i = 1$ (occupied) or $c_i = 0$ vacant. The complete list of these occupancies $\{c\}$ specifies a microstate. The average particle number density (fraction of occupied sites) is given by

$$c = L^{-d} \sum_i c_i$$

where L is the linear extent of the lattice and d its dimensionality.

The Hamiltonian of the lattice gas model is

$$H_{LG} = -\epsilon \sum_{\langle i,j \rangle} c_i c_j - \mu \sum_i c_i$$

where ϵ is an attraction energy between a pair of particles on adjacent (nearest neighbouring) sites and μ is a field known as the chemical potential, which couples to the particle density which is assumed to fluctuate around a mean value controlled by the prescribed chemical potential. This representation of the model in which the overall density fluctuates is known as the *Grand Canonical ensemble*.

6.1 Mapping between Ising model and lattice gas

The lattice gas model is interesting because whilst being a plausible model for a fluid, it maps onto the Ising model. This extends the applicability of the Ising model. To expose the mapping we write the grand partition function of the lattice gas:

$$\Xi = \sum_{\text{state}} \exp -\beta H_{LG} = \sum_{\{c\}} \exp \left[\beta \epsilon \sum_{\langle i,j \rangle} c_i c_j + \beta \mu \sum_i c_i \right]$$

where the sum is an unrestricted sum over the occupancies of the lattice sites. We now change variables to

$$c_i = (1 + s_i)/2; \quad J = \frac{\epsilon}{4} \quad h = \frac{\epsilon q + 2\mu}{4}$$

Hence

$$H_{LG} = H_1 + \text{constant}$$

Since the last term does not depend on the configuration, it feeds through as an additive constant in the free energy; and since all observables feature as derivatives of the free energy, the constant has no physical implications.

6.2 Phase diagram

Using these translation rules we can plot the phase diagram of the lattice gas in the number density-temperature plane.

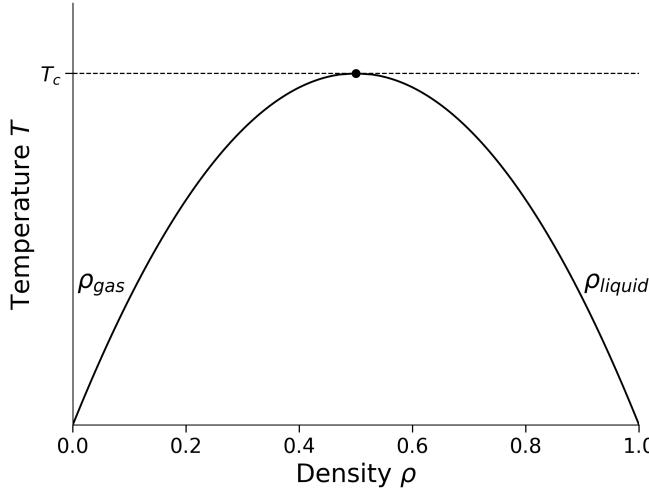


Figure 6.1: Phase diagram of the lattice gas model in the density-temperature plane.

In the $\mu-T$ plane there is a line of first order phase transitions terminating at a critical point. The first order line means that if $T < T_c$ we smoothly increase the chemical potential through the coexistence value of μ , the density of particles on our lattice $\rho = N/L^d$ jumps discontinuously from a low to a high value.

$$\rho_{\text{gas}} = \frac{1 - m^*}{2} \rightarrow \rho_{\text{liquid}} = \frac{1 + m^*}{2}$$

These values merge at T_c , the gas-liquid critical point. At higher temperatures, the distinction between the phases disappears.

6.3 Real Fluids

You may wish to compare this with the results of (say) van der Waals equation (see recommended textbooks for the required phase diagram). The main difference is that the lattice gas has so-called “particle-hole” symmetry, $\rho \rightarrow 1 - \rho$ (inherited from the up-down symmetry of the Ising model) which is not present for a real fluid. Accordingly, the phase diagram in a real fluid looks like a lopsided version of the above picture as shown in Figure ???. See [here](#) for some real experimental data showing the asymmetry of the coexistence curve in liquid metals.

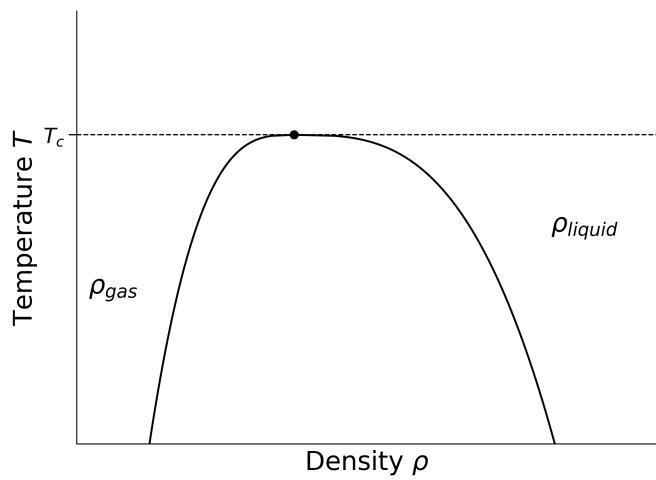


Figure 6.2: Schematic of the liquid-gas phase diagram in the $\rho - T$ plane for a realistic fluid .

Chapter 7

The Static Scaling Hypothesis

Historically, the first step towards properly elucidating near-critical behaviour was taken with the static scaling hypothesis. This is essentially a plausible conjecture concerning the origin of power law behaviour which appears to be consistent with observed phenomena. According to the hypothesis, the basis for power law behaviour (and associated scale invariance or “scaling”) 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 *generalized 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 generalized homogeneous function if it has the property

$$g(\lambda^a u, \lambda^b v) = \lambda g(u, v)$$

for all λ , where the parameters a and b (known as scaling parameters) are constants. An example of such a function is $g(u, v) = u^3 + v^2$ with $a = 1/3, b = 1/2$.

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

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

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

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

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 g nor the parameters a and b .

Now, the static scaling hypothesis asserts that in the critical region, the free energy F is a generalized homogeneous function of the (reduced) thermodynamic fields $t = (T - T_c)/T_c$ and $h = (H - H_c)$. Remaining with the example ferromagnet, the following scaling assumption can then be made:

$$F(\lambda^a t, \lambda^b h) = \lambda F(t, h).$$

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

Then

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

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

Now, as stated in Chapter ??, 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. Figure ??)

$$m(t, h) = -t^{(1-b)/a} m(1, t^{-b/a} h) \quad (7.1)$$

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

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

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

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

By taking further appropriate derivatives of the free energy, other relations between scaling parameters and critical exponents may be deduced. Such calculations (*Exercise: try to derive them*) yield the results $\delta = b/(1-b)$, $\gamma = (2b-1)/a$, and $\alpha = (2a-1)/a$. 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:-

$$\begin{aligned} \alpha + \beta(\delta + 1) &= 2 \\ \alpha + 2\beta + \gamma &= 2 \end{aligned}$$

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.

7.1 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 Equation ?? . Eliminating the scaling parameters a and b in favour of the exponents β and δ gives

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

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 He³ or Xe.

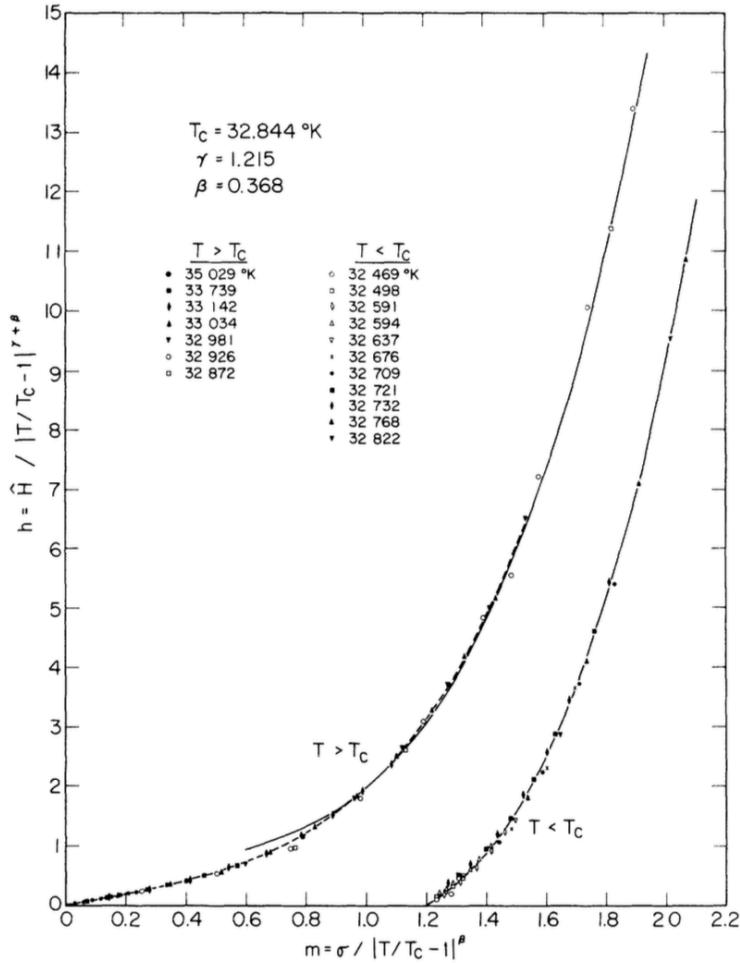


Figure 7.1: Magnetisation of CrBr₃ in the critical region plotted in scaled form (see text). From Ho and Lister (1969).

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 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.

7.2 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. Figure ?? shows a schematic example for the susceptibility of a magnet.

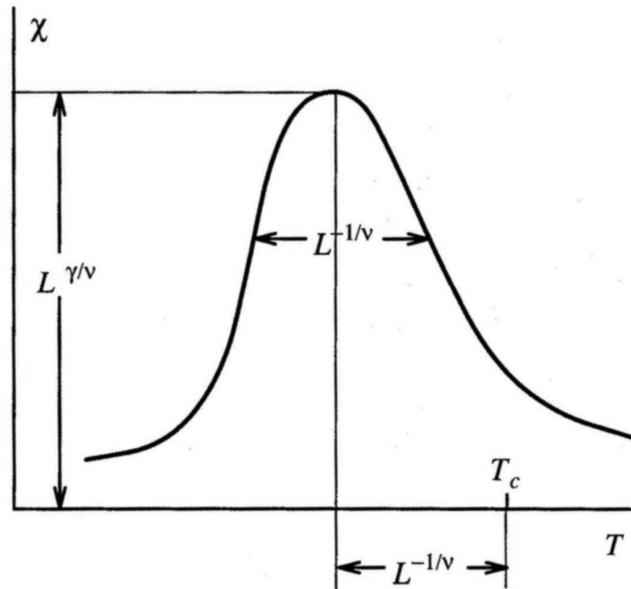


Figure 7.2: Schematic of the near-critical temperature dependence of the magnet susceptibility in a finite-sized system.

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 computational constraints prevent access to the largest system sizes which would provide accurate estimates 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 Section ??

Chapter 8

Universality and the Renormalisation Group Theory of Critical Phenomena

Owing to the absence of a wholly appropriate textbook for the material covered in this section, I have supplied more detailed notes than used in other parts of the unit.

The critical region is characterised by correlated microstructure on *all* length-scales up to and including the correlation length. Such a profusion of degrees of freedom can only be accurately characterized by a very large number of variables. 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, which stands today as the cornerstone of the modern theory of critical phenomena.

8.1 The critical point: A many length scale problem

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

1. The correlation length, ξ , ie the size of correlated microstructure.
2. 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.
3. Macroscopic size L_{\max} eg. size of the system.

The authentic critical region is defined by a window condition:

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

The physics of this regime is hard to tackle by analytic theory because it is characterized 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. Figure ??). 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.

8.2 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 characterize 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 Figure ???. 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:

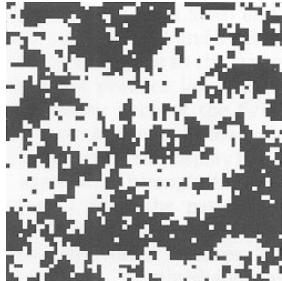
$$S_I(L) = L^{-d} \sum_i^I s_i \quad (8.1)$$

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 .

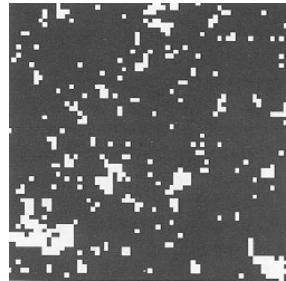
The coarse graining operation is easily implemented on a computer. In so doing one is faced with the fact that while the underlying Ising spins can only take two possible values, the block variables $S_I(L)$ have $L^d + 1$ possible values. Accordingly in displaying the consequences of the blocking procedure, we need a more elaborate colour convention than that used in Figure ???. 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 Figure ?? and Figure ???. 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.

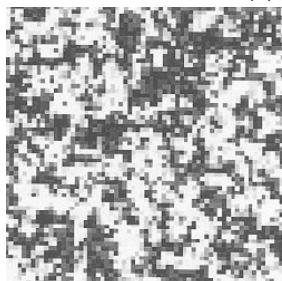
Consider first a system marginally above its critical point at a temperature T chosen so that the correlation length ξ is approximately 6 lattice spacing units. A typical arrangement (without coarse-graining) is shown in Figure ??(ai). The succeeding figures, Figure ??(aii) and Figure ??(aiii), show 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



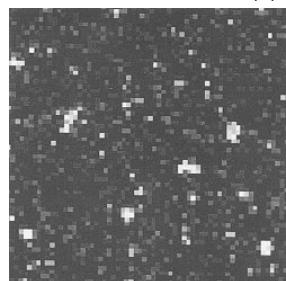
(a) (ai)



(b) (bi)



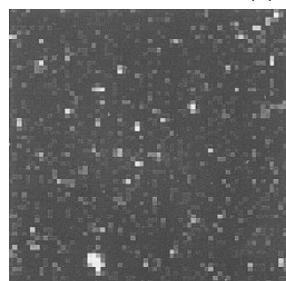
(c) (aii)



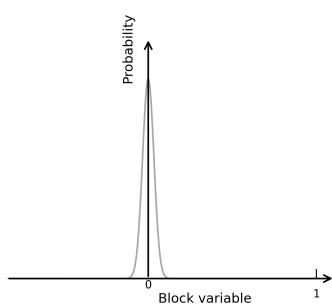
(d) (bii)



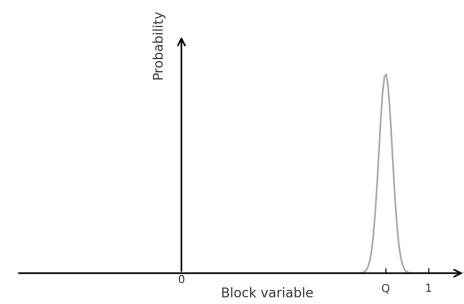
(e) (aiii)



(f) (biii)



(g) (aiv)



(h) (biv)

Figure 8.1: See text for details

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 Figure ??(aiii) and in an alternative form in Figure ??(aiv), which shows the limiting distribution of the coarse grained variables, averaged over many realizations 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. Figure ??(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 (Figure ??(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 (Figure ??(biv)).

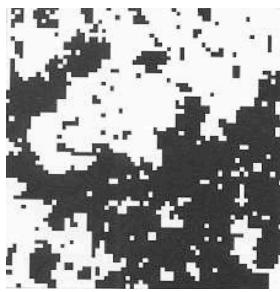
Consider now the situation *at* the critical point. Figure ??(ai) shows a typical arrangement; Figure ??(aii) and (aiii) show the results of coarse-graining with $L = 4$ and $L = 8$ respectively. Since the ξ is as large as the system itself the coarse graining does not produce less critical representations 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) shows (Figure ??(iv)) that it is almost identical to that of the $L = 8$ configurations (given the block variable scaling). 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 summarize. 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. A nice illustration of critical point scale invariance in the Ising model can be viewed [here](#).

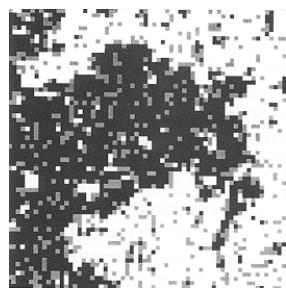
8.3 Universality and Scaling

Equipped with the coarse-graining technique, we now address the universality phenomenon. We aim to understand how it is that systems that are different microscopically can nevertheless display critical point behaviour which (in certain respects) is quantitatively identical.

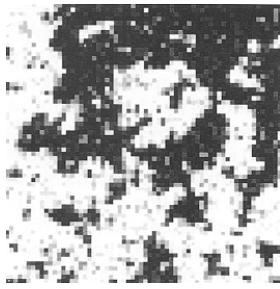
To obtain initial insight we introduce a spin-1 Ising model in which the spins take on three values ($s_i = 1, 0, -1$), in contrast to the two values ($s_i = 1, -1$) of the spin-1/2 Ising model. The two models have properties which are different: for example, T_c for the three-state model is some 30% lower than that of the two-state model (for the same coupling J). However, there is abundant evidence that the two models have the same universal properties.



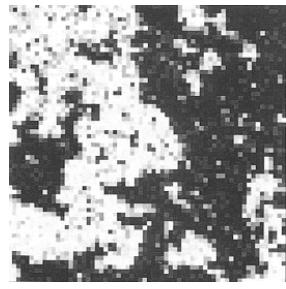
(a) (ai)



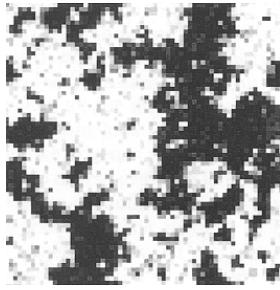
(b) (bi)



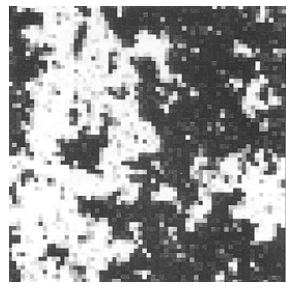
(c) (aii)



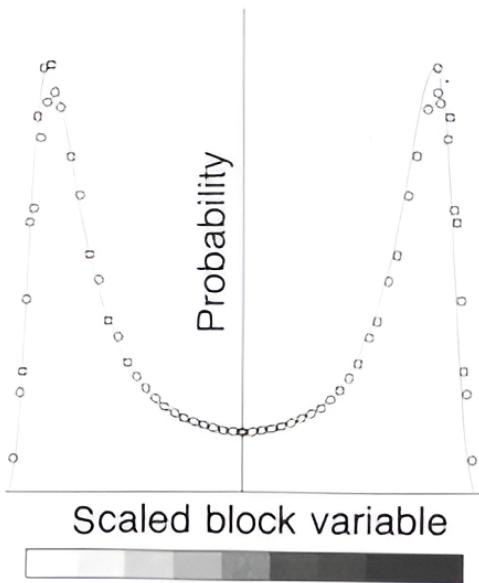
(d) (bii)



(e) (aiii)



(f) (biii)



(g) (iv)

Figure 8.2: See text for details

Let us explore what is the same and what is different in the configurations of the two models at criticality. The configurations of the local variables s_i are clearly qualitatively different for the two models. Now 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! This is expressed in Figure ??(iv), which shows the near coincidence of the distribution of block variables (grey-levels) for the two different coarse-graining lengths. Thus the coarse-graining erases the physical differences apparent in configurations where the local behaviour is resolvable, and exposes a profound configurational similarity.

8.3.1 Fluid-magnet universality

Let us now turn to fluid-magnet universality. 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 of side L) or more precisely, its fluctuation from its macroscopic average (Figure ??). 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 scales and the differences between the variable scales.

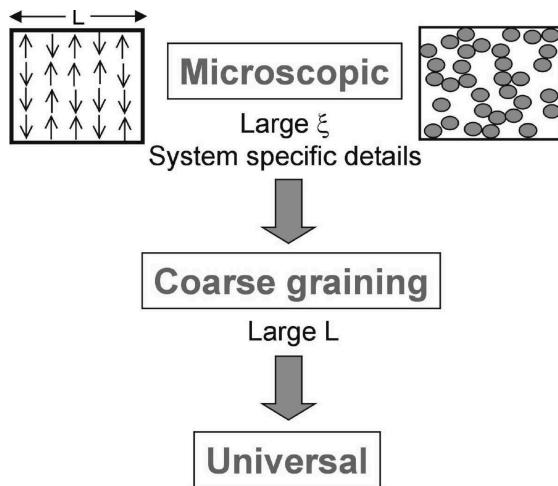


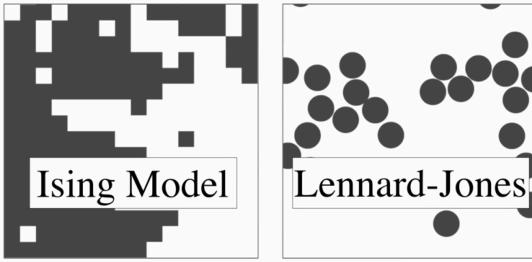
Figure 8.3: Schematic representation of the coarse graining operation via which the universal properties of fluids and magnets may be exposed.

The results is illustrated in Figure ??.

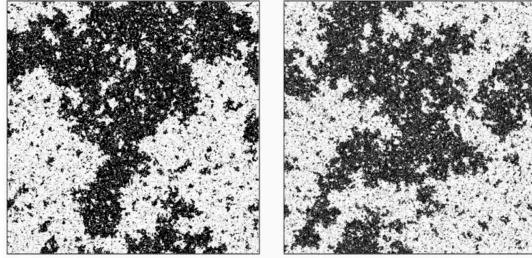
A movie in which we progressively zoom out shows how the loss of microscopic details reveals the large lengthscale universal features. [Movies/universality.mov](#)

8.4 Near critical scaling

The similarity of coarse-grained configurations of different systems is not restricted to the critical temperature itself. Suppose we have a two state spin model and a three state spin model each somewhat above their critical points at reduced temperature t . The two systems will have



(a) 2D critical Ising model and 2d critical Lennard-Jones fluid at small lengthscales



(b) Same models as above, but viewed at large lengthscales

Figure 8.4: 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.

somewhat different correlation lengths, ξ_1 and ξ_2 say. Suppose however, 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. Precisely what they look like depends upon our choice of ξ/L .

8.5 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 exponents. 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.

| | $d = 2$ | $d = 3$ |
|-----------------------------------|---------------|-------------------|
| Critical temperature | 0.5673 | 0.75 |
| Order parameter exponent β | $\frac{1}{8}$ | 0.325 ± 0.001 |
| Susceptibility exponent γ | $\frac{7}{4}$ | 1.24 ± 0.001 |
| Correlation length exponent ν | 1 | 0.63 ± 0.001 |

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 $n = 2$ (the so-called XY model) or $n = 3$ (Heisenberg model), 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.

8.6 Critical exponents

We consider now how the critical exponents, may be computed via the coarse-graining procedure. 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 t is sufficiently small that $\xi \gg L_{\min}$. Universality and scaling may be expressed in the claim that, for *any* L and t , scale factors $a(L)$ and $b(L)$ may be found such that the probability distribution $p(S(L), t)$ can be written in the form

$$p(S(L), t) = b(L)\tilde{p}(b(L)S(L), a(L)t) \quad (8.2)$$

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 Section ???. The critical exponents are implicit in the L -dependence of these scale factors. Specifically one finds:

$$\begin{aligned} a(L) &= a_0 L^{1/\nu} \\ b(L) &= b_0 L^{\beta/\nu} \end{aligned} \quad (8.3)$$

where the amplitudes a_0 and b_0 are system specific (non-universal) constants.

These results state that the critical exponents (in the form $1/\nu$ and β/ν) characterize the ways in which the configuration spectrum evolves under coarse-graining. Consider, first the exponent ratio 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-graining length L_2 only if the variable scale in the latter configurations is amplified. The required amplification follows from Equation ?? and Equation ???: it is

$$\frac{b(L_2)}{b(L_1)} = \left(\frac{L_2}{L_1} \right)^{\beta/\nu}.$$

The exponent ratio β/ν thus controls the rate at which the scale of the ordering variable decays with increasing coarse-graining length.

Consider now the exponent $1/\nu$. 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 appearance. More precisely, we see from Equation ?? 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

$$\frac{a(L_2)}{a(L_1)} = \left(\frac{L_2}{L_1}\right)^{1/\nu}.$$

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 exponent $1/\nu$ controls the rate at which the effective reduced temperature flows with increasing coarse-graining length.

8.7 Finite-size scaling

We can exploit the fact that the scale factors $a(L)$ and $b(L)$ depend on critical exponents to estimate the values of these exponents using computer simulation. Consider the average of the block variable $S(L)$. Consideration of Equation ?? shows that this is non other than the value of the order parameter Q , measured over a block of side L . Thus from the definition of an average

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

where $p(S(L))$ is the probability distribution of $S(L)$.

Making use of the representation of Equation ??, we then have that

$$Q(L, t) = \int b(L)S(L)\tilde{p}(b(L)S(L), a(L)t)dS(L)$$

To integrate this we need to change the integration variable from $S(L)$ to $b(L)S(L)$. We have $d(b(L)S(L)) = b(L)dS(L)$ since $b(L)$ does not fluctuate. Thus

$$\begin{aligned} Q(L, t) &= b^{-1}(L) \int b(L)S(L)\tilde{p}(b(L)S(L), a(L)t)d(b(L)S(L)) \\ &= b^{-1}(L)f(a(L)t) \\ &= b_0 L^{-\beta/\nu} f(a_0 L^{1/\nu} t) \end{aligned}$$

where f is a universal function (defined as the first moment of $\tilde{p}(x, y)$ with respect to y).

The above results provide a method 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. Section ??). However, we know that its value must vanish in the limit of infinite L ; it does so like

$$Q(L, 0) = b_0 L^{-\beta/\nu} f(0) \equiv Q_0 L^{-\beta/\nu}$$

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 $1/\nu$.

8.8 Summary of main points

As this is quite a long chapter let us summarise the main points:

- **Limitations of conventional theories:** Mean field theories and the scaling hypothesis are insufficient in the critical region due to their neglect of correlations across all relevant length scales.
- **Critical region characteristics:** Near-critical systems exhibit correlated microstructure on all length scales up to the correlation length. The complexity of this structure makes both analytical and computational study challenging.
- **Relevant length scales:**
 1. Correlation length (ξ)
 2. Minimum microscopic scale (L_{\min})
 3. Macroscopic system size (L_{\max})
- **Window Condition for criticality:** The true critical regime satisfies $L_{\max} \gg \xi \gg L_{\min}$, encompassing a broad range of scales where complex, often fractal-like, structures are present.
- **Renormalisation Group (RG) methodology:** RG involves the stepwise elimination of degrees of freedom by coarse-graining the system over increasing length scales. A fourth scale, L , represents the resolution at which the system is described.
- **Effect of coarse-graining:** Coarse-graining changes the effective reduced temperature, captured by the relation $a(L)t$, where $a(L)$ scales with L as $L^{1/\nu}$. This helps describe how critical configurations evolve with resolution.
- **Universality:**
 - Coarse-graining reveals that microscopically different systems can exhibit identical critical behavior when observed at large scales.
 - The concept of universality explains why disparate systems, such as magnets and fluids, can show the same critical exponents and scaling laws.
 - Critical behavior depends primarily on general features like dimensionality and symmetry, rather than microscopic details.
- **Finite-Size scaling:**
 - The average block variable $Q(L, t)$ is related to block size L and reduced temperature t through scaling relations. Computer simulations exploit this to extract scaling functions and the values of critical exponents.

8.9 Addendum: The effective coupling viewpoint of the renormalization group (non examinable)

🔥 Notes for those interested in a different perspective on RG theory.

Let us begin by returning to our fundamental Equation ??, which we rewrite as

$$p = Z^{-1} e^{-H}$$

where $H \equiv E/k_B T$.

The first step is then to imagine that we generate, by a computer simulation procedure for example, a sequence of configurations with relative probability $\exp(-H)$. 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 H' of the coarse-grained variables which would produce these coarse-grained configurations with the correct relative probability $\exp(-H')$? Clearly the form of H' depends on the form of H thus we can write symbolically

$$H' = R(H)$$

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 (RG). 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 utilizing the configurational energy and appealing to the basic machinery of statistical mechanics set out in Equation ?? and Equation ??, 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 coarse-grained system described by the energy H' . The large-distance properties of this system are the same as the large-distance properties of the physical system, since coarse-graining operation preserves large-scale configurational structure. In this representation the problem is a little easier: while the ξ associated with H' is the same as the ξ associated with H , the minimum length scale of H' is bigger than that of H . Thus the statistical mechanics of 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 to solve. The benefits accruing from this procedure may be amplified by repeating it. Repeated application of R will eventually result in a coarse-grained energy function describing configurations in which the ξ is no bigger than the minimum length scale. The associated system is far from criticality and its properties may be reliably computed by any of a wide variety of approximation schemes. These properties are the desired large-distance properties of the physical system. 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 setup the framework for a simple illustrative example, let us return to the lattice Ising model 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, \dots , or $J_a, a = 1, 2, \dots$. In reduced units, the equivalent coupling strengths are $K_a = J_a/k_B T$. Their values determine uniquely the energy for any given configuration.

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 H the transformation will in general produce others in H' .

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. The new energy function 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

$$K'_a = f_a(K_1, K_2, \dots) = f_a(\mathbf{K}), \quad a = 1, 2, \dots \quad (8.4)$$

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

8.9.1 A simple example

This example illustrates how one can perform the RG transformation Equation ?? 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.

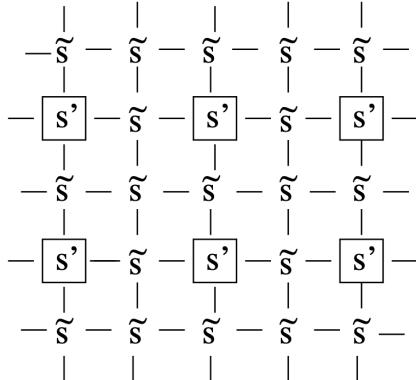


Figure 8.5: Coarse graining by decimation. The spins on the original lattice are divided into two sets $\{s'\}$ and $\{\tilde{s}\}$. The $\{s'\}$ spins occupy a lattice whose spacing is twice that of the original. The effective coupling interaction between the $\{s'\}$ spins is obtained by performing the configurational average over the $\{\tilde{s}\}$

Consider an Ising model in two dimensions, with only nearest neighbour interactions as shown in Figure ???. 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 H' for the s' spins by performing an average over all the possible arrangements of the \tilde{s} spins

$$\exp(-H') = \sum_{\{\tilde{s}\}} \exp(-H). \quad (8.5)$$

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 realizes two basic aims of the RG method: the long-distance physics of the ‘original’ system, described by H , is contained in that of the ‘new’ system, described by H' (indeed the partition functions are the same as one can see by summing both sides over s') and the new system is further from critically because the ratio of ξ 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 Equation ???. 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. In its simplest mathematical form, since H contains a factor $1/k_B T$, it involves the expansion of $\exp(-H)$ as a power series:

$$\exp(-H/k_B T) = 1 - H/k_B T + \frac{1}{2!}(H/k_B T)^2 + \dots$$

We substitute this expansion into the right hand side of Equation ?? 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 (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 ($H^2/2$) does contribute. If one writes out explicitly the form of $H^2/2$ one finds terms of the form $K^2 s'_1 \tilde{s} s'_2 = K^2 s'_1 s'_2$, 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 Equation ??, 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

$$K' = K^2 \quad (8.6)$$

Of course many other terms and couplings are generated by the higher orders of the high temperature expansion and it is necessary to include these if one wishes reliable values for the critical temperature and exponents. However, our aim here is to use this simple calculation to illustrate the RG method. Let us therefore close our eyes, forget about the higher order terms and show how the RGT Equation ?? can be used to obtain information on the phase transition.

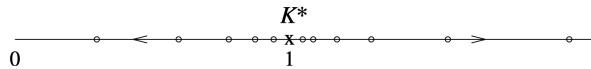


Figure 8.6: Coupling flow under the decimation transformation described in the text.

The first point to note is that that mathematically Equation ?? 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 Equation ?? is away from the fixed point, as illustrated in Figure ??.

The physical significance of this is as follows: suppose that the original system is at its critical point so that the ratio of ξ 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 Equation ??, 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 criticality, 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 ξ is finite and the new system is further from critically because the ratio of ξ 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

$$K_c = J/k_B T_c = 1 \quad (8.7)$$

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 (Equation ??) in terms of the deviation of the coupling from its fixed point value. A Taylor expansion of the function $K' = K^2$ yields

$$\begin{aligned} K' &= (K^*)^2 + (K - K^*) \left. \frac{\partial K'}{\partial K} \right|_{K=K^*} + \frac{1}{2} (K - K^*)^2 \left. \frac{\partial^2 K'}{\partial K^2} \right|_{K=K^*} + \dots \\ K' - K^* &= 2(K - K^*) + (K - K^*)^2 \end{aligned}$$

where in the second line we have used the fact that the first derivative evaluates to $2K^* = 2$ and $(K^*)^2 = K^*$.

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:

$$t' = 2t$$

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

$$\xi' = \xi/2$$

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

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

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- the prediction in Eq. refeq:Kc 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, Equation ??, 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.

8.9.2 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.

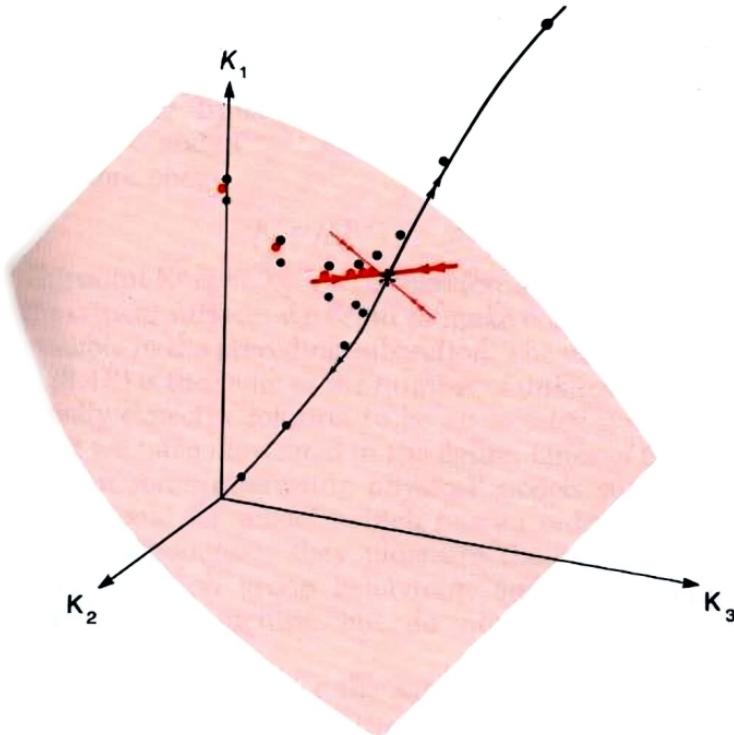


Figure 8.7: General flow in coupling space

Figure ?? is a representation of the space of all coupling strengths K_a in the energy function $H/k_B T$. 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, \dots . As the temperature T varies, the coupling strengths $K_a = J_a/k_B T$ trace out a straight line, or ray, from the origin of the space in the direction (J_1, J_2, J_3, \dots) . 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, $K_a = J_a/k_B T, a = 1, 2, \dots$. 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 ξ . It is shown schematically as the shaded surface in Figure ??.

(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 temperature 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 Figure ??.

system, with coupling strengths $K_a, a = 1, 2, \dots$. 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: $\mathbf{K} \rightarrow \mathbf{K}^{(1)} \rightarrow \dots \rightarrow \mathbf{K}^{(n)}$ 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 ξ to coarse-graining scale is reduced by a factor of b^{-n} . The dots in Figure ?? 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 ξ 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, $\mathbf{K}^{(1)}, \mathbf{K}^{(2)}, \dots, \mathbf{K}^{(n)}$, 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

$$K_a^* = f_a(K^*) \quad a = 1, 2, \dots$$

This point $\mathbf{K}^* \equiv K_1^*, K_2^*, \dots$ 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 $K_a^{(1)}, K_a^{(2)}, \dots$ 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 Section ?? . 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 Figure ?? 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.

Chapter 9

The renormalisation group: effective coupling viewpoint

Let us begin by returning to our fundamental Equation ??, which we rewrite as

$$p = Z^{-1} e^{-H}$$

where $H \equiv E/k_B T$.

The first step is then to imagine that we generate, by a computer simulation procedure for example, a sequence of configurations with relative probability $\exp(-H)$. 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 H' of the coarse-grained variables which would produce these coarse-grained configurations with the correct relative probability $\exp(-H')$? Clearly the form of H' depends on the form of H thus we can write symbolically

$$H' = R(H)$$

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 utilizing the configurational energy and appealing to the basic machinery of statistical mechanics set out in Equation ?? and Equation ??, 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 coarse-grained system described by the energy H' . The large-distance properties of this system are the same as the large-distance properties of the physical system, since coarse-graining operation preserves large-scale configurational structure. In this representation the problem is a little easier: while the ξ associated with H' is the same as the ξ associated with H , the minimum length scale of H' is bigger than that of H . Thus the statistical mechanics of 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 to solve. The benefits accruing from this procedure may be amplified by repeating it. Repeated application of R will eventually result in a coarse-grained energy function describing configurations in which the ξ is no bigger than the minimum length scale. The associated system is far

from criticality and its properties may be reliably computed by any of a wide variety of approximation schemes. These properties are the desired large-distance properties of the physical system. 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 setup the framework for a simple illustrative example, let us return to the lattice Ising model 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, \dots , or $J_a, a = 1, 2, \dots$. In reduced units, the equivalent coupling strengths are $K_a = J_a/k_B T$. Their values determine uniquely the energy for any given configuration.

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 H the transformation will in general produce others in H' .

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. The new energy function 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

$$K'_a = f_a(K_1, K_2, \dots) = f_a(\mathbf{K}), \quad a = 1, 2, \dots \quad (9.1)$$

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

9.1 A simple example

This example illustrates how one can perform the RG transformation Equation ?? 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.

Consider an Ising model in two dimensions, with only nearest neighbour interactions as shown in Figure ???. 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 H' for the s' spins by performing an average over all the possible arrangements of the \tilde{s} spins

$$\exp(-H') = \sum_{\{\tilde{s}\}} \exp(-H). \quad (9.2)$$

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 realizes two basic aims of the RG method: the long-distance physics of the ‘original’ system, described by H , is contained in that of the ‘new’ system, described by H' (indeed the partition functions are the same as one can see by summing both sides over s') and the new system is further from criticality because the ratio of ξ to lattice spacing (‘minimum length

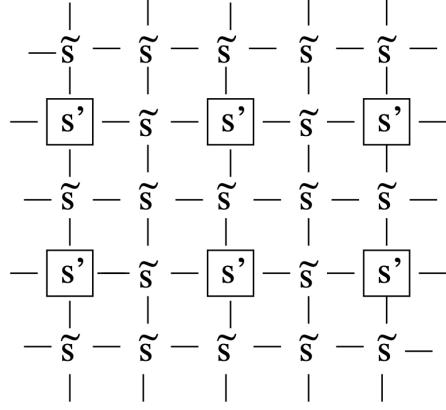


Figure 9.1: Coarse graining by decimation. The spins on the original lattice are divided into two sets $\{s'\}$ and $\{\tilde{s}\}$. The $\{s'\}$ spins occupy a lattice whose spacing is twice that of the original. The effective coupling interaction between the $\{s'\}$ spins is obtained by performing the configurational average over the $\{\tilde{s}\}$

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 Equation ???. 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. In its simplest mathematical form, since H contains a factor $1/k_B T$, it involves the expansion of $\exp(-H)$ as a power series:

$$\exp(-H/k_B T) = 1 - H/k_B T + \frac{1}{2!}(H/k_B T)^2 + \dots$$

We substitute this expansion into the right hand side of Equation ?? 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 (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 ($H^2/2$) does contribute. If one writes out explicitly the form of $H^2/2$ one finds terms of the form $K^2 s'_1 \tilde{s} s'_2 = K^2 s'_1 s'_2$, 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 Equation ??, 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

$$K' = K^2 \tag{9.3}$$

Of course many other terms and couplings are generated by the higher orders of the high temperature expansion and it is necessary to include these if one wishes reliable values for the critical temperature and exponents. However, our aim here is to use this simple calculation to illustrate the RG method. Let us therefore close our eyes, forget about the higher order terms and show how the RGT Equation ?? can be used to obtain information on the phase transition.

The first point to note is that that mathematically Equation ?? 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

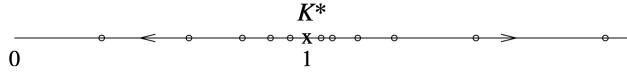


Figure 9.2: Coupling flow under the decimation transformation described in the text.

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 Equation ?? is away from the fixed point, as illustrated in Figure ???. The physical significance of this is as follows: suppose that the original system is at its critical point so that the ratio of ξ 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 Equation ??, 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 ξ is finite and the new system is further from critically because the ratio of ξ 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

$$K_c = J/k_B T_c = 1 \quad (9.4)$$

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 (Equation ??) in terms of the deviation of the coupling from its fixed point value. A Taylor expansion of the function $K' = K^2$ yields

$$\begin{aligned} K' &= (K^*)^2 + (K - K^*) \frac{\partial K'}{\partial K} \Big|_{K=K^*} + \frac{1}{2}(K - K^*)^2 \frac{\partial^2 K'}{\partial K^2} \Big|_{K=K^*} + \dots \\ K' - K^* &= 2(K - K^*) + (K - K^*)^2 \end{aligned}$$

where in the second line we have used the fact that the first derivative evaluates to $2K^* = 2$ and $(K^*)^2 = K^*$.

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:

$$t' = 2t$$

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

$$\xi' = \xi/2$$

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

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

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- the prediction in Eq. refeq:Kc for K_c 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, Equation ??, 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.

9.2 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.

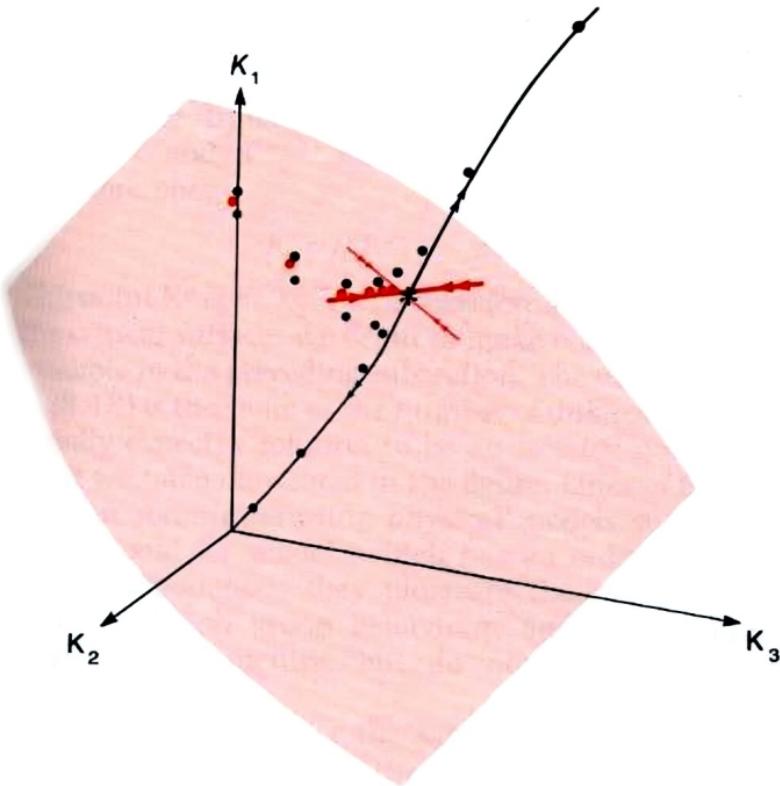


Figure 9.3: General flow in coupling space

Figure ?? is a representation of the space of all coupling strengths K_a in the energy function $H/k_B T$. 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, \dots . As the temperature T varies, the coupling strengths $K_a = J_a/k_B T$ trace out a straight line, or ray, from the origin of the space in the direction (J_1, J_2, J_3, \dots) . 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, $K_a = J_a/k_B T, a = 1, 2, \dots$. 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 ξ . It is shown schematically as the shaded surface in Figure ???. (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 temperature 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 Figure ???. Suppose we start with a physical system, with coupling strengths $K_a, a = 1, 2, \dots$. 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: $\mathbf{K} \rightarrow \mathbf{K}^{(1)} \rightarrow \dots \rightarrow \mathbf{K}^{(n)}$ 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 ξ to coarse-graining scale is reduced by a factor of b^{-n} . The dots in Figure ?? 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 ξ 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, $\mathbf{K}^{(1)}, \mathbf{K}^{(2)}, \dots, \mathbf{K}^{(n)}$, 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

$$K_a^* = f_a(K^*) \quad a = 1, 2, \dots$$

This point $\mathbf{K}^* \equiv K_1^*, K_2^*, \dots$ 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 $K_a^{(1)}, K_a^{(2)}, \dots$ 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 Section ???. 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 Figure ???) 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.

Chapter 10

Dynamics of first order phase transitions: nucleation, growth and spinodal decomposition

10.1 Introduction to nucleation

In previous discussions, we considered first-order phase transitions but deferred a detailed analysis of the *dynamical* mechanism by which a system evolves from one phase to another. We now address this question explicitly.

Consider again the Ising model at a temperature $T < T_c$, where the system is initially prepared in the majority spin-up phase at zero external field, $H = 0$. We now examine the system's response to the application of a small negative external field $H < 0$, which lowers the free energy of the spin-down phase relative to the spin-up phase.

Despite the global free energy favoring the spin-down phase, the system does not undergo an instantaneous transition. This delay is a consequence of the free energy barrier associated with nucleating a region of the stable phase within the metastable one, as introduced in Section ???. The dynamical pathway of the phase transition proceeds via *nucleation* of localized regions—referred to as *droplets*—of the stable (spin-down) phase embedded within the metastable (spin-up) background. Once nucleated, these droplets may grow over time and ultimately coalesce to transform the system to the stable phase.

The nucleation of a droplet of the stable phase of size n spins entails a competition between bulk and interfacial contributions to the free energy. The bulk free energy gain is linear in n , given by $-nH$, due to the alignment of spins with the external field. However, this gain is offset by an interfacial free energy cost arising from broken bonds at the boundary between phases. For the Ising model, each broken bond contributes an energy cost of $+2J$, so the total interfacial energy scales with the perimeter (in 2D) or surface area (in 3D) of the droplet. This interfacial contribution is referred to as the *surface tension*, and constitutes a true free energy cost: it includes not only the energetic penalty from broken bonds but also an entropic contribution due to the configurational degrees of freedom associated with the droplet shape.

The resulting competition between the extensive free energy gain and the sub-extensive interfacial cost leads to a free energy barrier for droplet formation. Only fluctuations that produce a droplet larger than a critical size n_c will grow; smaller droplets will shrink. This framework is formalized in *classical nucleation theory*, which provides a quantitative description of the nucle-

ation rate, critical droplet size, and the associated activation energy barrier.

10.2 Classical Nucleation Theory: Homogeneous Nucleation

We now present the framework of **classical nucleation theory** (CNT) for the case of *homogeneous nucleation*, in which the nucleation of the stable phase occurs spontaneously and uniformly throughout the bulk of the metastable phase, without the aid of impurities, defects, or surfaces.

Let us consider a droplet of the stable (spin-down) phase of radius R embedded within the metastable (spin-up) background. The total change in free energy $\Delta F(R)$ associated with forming such a droplet consists of two competing contributions:

1. **Bulk free energy gain:** The interior of the droplet consists of $V \sim R^d$ spins aligned with the external field $H < 0$, leading to a volume free energy change

$$\Delta F_{\text{bulk}}(R) = -|\Delta f| R^d,$$

where $|\Delta f| \propto |H|$ is the free energy density difference between the metastable and stable phases, and d is the spatial dimensionality of the system.

2. **Interfacial free energy cost:** The boundary between the two phases has a surface area scaling as R^{d-1} , and incurs a free energy cost proportional to the *surface tension* σ :

$$\Delta F_{\text{surface}}(R) = \sigma S_d R^{d-1},$$

where S_d is a geometrical factor (e.g., $S_2 = 2\pi$ in 2D and $S_3 = 4\pi$ in 3D).

The total free energy change is therefore given by

$$\Delta F(R) = \sigma S_d R^{d-1} - |\Delta f| V_d R^d,$$

where V_d is another dimension-dependent constant. This expression (Figure ??) exhibits a characteristic maximum at a critical droplet radius R_c , obtained by extremizing $\Delta F(R)$ with respect to R :

$$\frac{d\Delta F}{dR} = 0 \quad \Rightarrow \quad R_c = \frac{(d-1)\sigma S_d}{d|\Delta f| V_d}.$$

The corresponding free energy barrier for nucleation is

$$\Delta F_c = \Delta F(R_c) = \frac{(d-1)^{d-1}}{d^d} \cdot \frac{(S_d)^d \sigma^d}{(|\Delta f|)^{d-1} (V_d)^{d-1}}.$$

This barrier must be surmounted by thermal fluctuations in order for a critical nucleus to form and grow. The nucleation rate per unit volume is given (in the Arrhenius approximation) by

$$I \sim I_0 \exp \left(-\frac{\Delta F_c}{k_B T} \right),$$

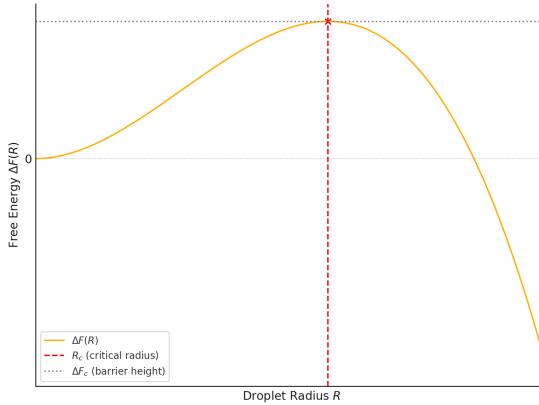


Figure 10.1: Free energy barrier $\Delta F(r)$ for nucleation of a spherical droplet as a function of radius R (schematic)

where I_0 is a prefactor determined by microscopic kinetics, and k_B is Boltzmann's constant.

10.2.1 Interpretation and Scaling Behavior

Several key features emerge from this analysis:

- **Barrier scaling:** The nucleation barrier $\Delta F_c \sim \sigma^d / |\Delta f|^{d-1}$ diverges as $H \rightarrow 0$, reflecting the increasing stability of the metastable phase near the coexistence point.
- **Critical radius:** The critical droplet size $R_c \sim \sigma / |\Delta f|$ also diverges as $|\Delta f| \rightarrow 0$, indicating that larger fluctuations are required to initiate nucleation close to the coexistence line.
- **Dimensional dependence:** Both ΔF_c and R_c exhibit strong dependence on the spatial dimension d , with nucleation becoming increasingly suppressed in higher dimensions due to the dominance of interfacial cost.

In summary, homogeneous nucleation in a first-order transition is governed by a delicate balance between surface tension and bulk free energy gain. Only droplets exceeding a critical size can overcome the barrier and initiate a transition. This sets an intrinsic timescale for the dynamics of phase transformation, which can become extremely long near coexistence due to the exponentially small nucleation rate.

10.3 Domain growth

Following successful nucleation of a supercritical droplet, the system enters a regime where the global transformation is driven by the deterministic growth of domains of the stable phase. Such *domain growth* involves more atoms or molecules attaching to these nuclei, causing them to expand into larger structures (e.g., growing crystals or droplets).

Growth rate depends on factors like temperature, concentration, and the availability of building blocks.

The shape and structure of the final phase often depend on how growth occurs (e.g., slow growth may form perfect crystals; rapid growth may be irregular).

10.4 Spinodal decomposition

Previously we have considered a scenario in which we move our system to a statepoint just inside the coexistence region so that the original phase remains metastable. We then wait for a fluctuation that yields a critical nucleus and subsequent growth.

Now imagine that rather than positioning the system just inside the coexistence region we move immediately to a state point well inside the coexistence region (cf. Figure ??) such that there is no metastable minimum in the free energy. Then there is no nucleation and growth, rather the system is *unstable* and immediately starts to phase separate at all points in the system. This so called *spinodal decomposition*.

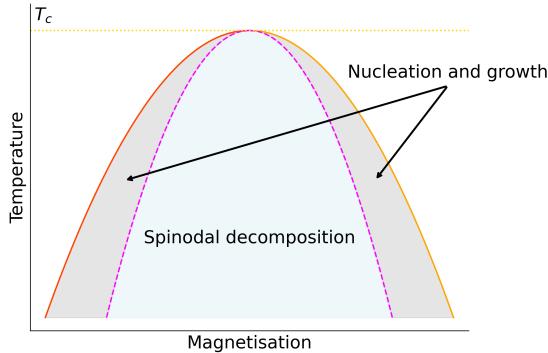


Figure 10.2: Schematic phase doagram showing the regions in which one observes nucleation and growth, or spinodal decomposition

The nature of spinodal decomposition and late-time domain growth depends sensitively on whether the order parameter is conserved or not.

10.4.1 Non-Conserved Order Parameter Dynamics (Model A)

In systems with a non-conserved order parameter, such as the Ising model with single spin flip (so called Glauber) dynamics, the order parameter can relax locally without constraint. This leads to curvature-driven motion of interfaces.

The typical domain size $L(t)$ grows algebraically with time:

$$L(t) \sim t^{1/2},$$

corresponding to a dynamic exponent $z = 2$. The growth is driven by reduction of the total interfacial area—regions of high curvature (small domains) shrink and are absorbed by larger, flatter ones.

Let us assume that domain walls move with velocity proportional to curvature:

$$v \sim \frac{1}{L}$$

Then with the typical domain size being $L(t)$ it follows that

$$\frac{dL}{dt} \sim \frac{1}{L}$$

Integrating both sides yields the $t^{1/2}$ domain growth law.

It turns out that the detailed evolution of the coarse-grained order parameter field $\phi(\mathbf{r}, t)$ is governed by the **Allen–Cahn equation**:

$$\frac{\partial \phi}{\partial t} = -\frac{\delta F[\phi]}{\delta \phi},$$

where $F[\phi]$ is a coarse-grained Ginzburg–Landau free energy functional:

$$F[\phi] = \int d^d x \left[\frac{1}{2} (\nabla \phi)^2 + V(\phi) \right],$$

with $V(\phi)$ typically a double-well potential such as $V(\phi) = \frac{1}{4}(\phi^2 - 1)^2$.

10.4.2 Conserved Order Parameter Dynamics (Model B)

For systems with a conserved order parameter, such as phase separation in binary alloys or the Ising model with spin-swap (so called ‘Kawasaki’) dynamics, the order parameter (e.g., composition or particle number) must be conserved locally. This imposes a diffusive constraint on the dynamics.

The domain size again grows algebraically, but with a different exponent:

$$L(t) \sim t^{1/3},$$

corresponding to a dynamic exponent $z = 3$.

A chemical potential difference drives diffusion and is given by

$$\Delta \mu \sim \frac{\sigma}{L}$$

(again due to curvature, where σ is surface tension)

Now the flux is proportional to the chemical potential gradient (Fick’s law):

$$\text{Flux} \sim -\nabla \mu \sim \frac{\Delta \mu}{L} \sim \frac{\sigma}{L^2}$$

and the rate of change of domain size is proportional to this flux:

$$\frac{dL}{dt} \sim \frac{1}{L^2} \quad \Rightarrow \quad L(t) \sim t^{1/3}$$

It turns out that the detailed dynamics are described by the **Cahn–Hilliard equation**, a continuity equation of the form:

$$\frac{\partial \phi}{\partial t} = \nabla^2 \left(\frac{\delta F[\phi]}{\delta \phi} \right),$$

reflecting that the order parameter can only evolve via diffusion of its conjugate chemical potential. This leads to the slow transport of material across domains and a more sluggish coarsening process compared to the non-conserved case.

10.4.3 Schematic of Domain Growth in 2D Ising model

Here are schematic illustrations of domain growth for:

- Non-conserved dynamics (Model A): Domains coarsen rapidly, with smoother and larger regions due to free relaxation of the order parameter.
- Conserved dynamics (Model B): Coarsening is slower and domains are more intricate, reflecting the constraint of local conservation.

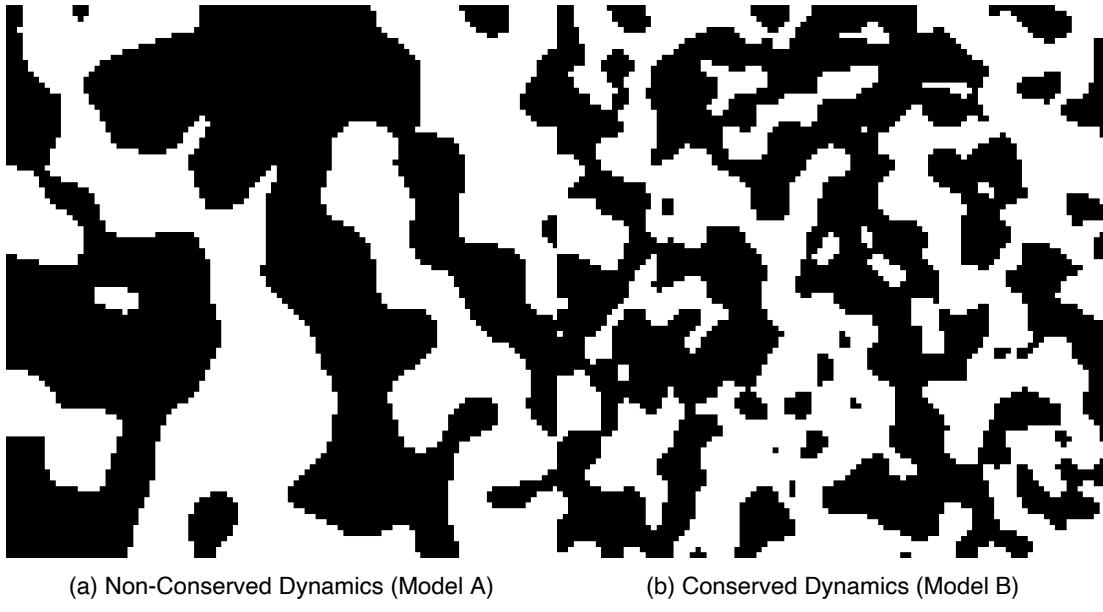


Figure 10.3: Schematic illustrations of domain morphology resulting from spinodal decomposition or late-time domain growth for non-conserved and conserved dynamics.

10.4.4 Dynamics Scaling Hypothesis

At late times, both conserved and non-conserved systems exhibit dynamic scaling: the statistical properties of the domain morphology become self-similar under rescaling of lengths by $L(t)$.

For example, the equal-time two-point correlation function satisfies

$$C(r, t) = f\left(\frac{r}{L(t)}\right),$$

where $f(x)$ is a time-independent scaling function. Plots of $C(r, t)$ collapse when plotted as a function of $r/L(t)$.

The structure factor $S(k, t)$, which is the Fourier transform of the correlation function, is experimentally accessible eg via X-ray or neutron scattering, and also obeys dynamic scaling:

$$S(k, t) = \int d^d r e^{-i\vec{k}\cdot\vec{r}} C(r, t).$$

Substituting the scaling form of $C(r, t)$ into this expression, and changing variables to $\vec{u} = \vec{r}/L(t)$, gives:

$$S(k, t) = L(t)^d \int d^d u e^{-i\vec{k}\cdot L(t)\vec{u}} f(u) = L(t)^d g(kL(t)),$$

with $g(x)$ a universal scaling function dependent on the dynamical class and dimensionality.

10.4.5 Summary of Growth Laws

| Dynamics Type | Conservation Equation Type | Growth Law | Dynamic Exponent | |
|----------------------------|----------------------------|---------------|---------------------|---------|
| Model A (e.g. Glauber) | No | Allen–Cahn | $L(t) \sim t^{1/2}$ | $z = 2$ |
| Model B (e.g. Kawasaki) | Yes | Cahn–Hilliard | $L(t) \sim t^{1/3}$ | $z = 3$ |

Remarks: The domain growth exponents $1/z$ are robust under many conditions, but can be modified in the presence of disorder, long-range interactions, or hydrodynamic effects.

In both the model A and model B cases, the system coarsens until it reaches equilibrium, characterized by a uniform macroscopic phase and the complete elimination of interfaces.

The approach to equilibrium is algebraically slow (described by power laws) due to the scale-free nature of domain dynamics.

More about the scientists mentioned in this chapter:

[John Cahn](#)

Chapter 11

Introduction to stochastic processes

Many natural phenomena are stochastic — they involve randomness in their evolution over time.

In classical physics, this randomness may arise from our lack of knowledge about microscopic details. For example, in a gas, we do not know the precise positions and velocities of each particle, so their collisions with the container walls appear random.

In quantum mechanics, stochasticity is even more intrinsic. The fundamental objects are probability amplitudes, and outcomes are inherently probabilistic.

To describe these systems effectively, we use a coarse-grained probabilistic description, which tracks the likelihood of different outcomes rather than precise trajectories. One of the key tools in this approach is the master equation.

11.1 The Master Equation

Consider a system in microstate i with energy E_i . It can transition to neighboring microstates j , where the energy difference $|E_j - E_i|$ is small (within δE).

Let ν_{ij} be the **rate** at which the system jumps from state i to state j . Over an infinitesimal time interval dt , the probability p_i changes as:

$$dp_i = \left[-p_i \sum_j \nu_{ij} + \sum_j \nu_{ji} p_j \right] dt$$

This expression contains two terms:

- A *loss term*: the system leaves state i at rate ν_{ij} ,
- A *gain term*: the system arrives in state i from other states j at rate ν_{ji} .

The master equation becomes:

$$\frac{dp_i}{dt} = - \sum_j \nu_{ij} p_i + \sum_j \nu_{ji} p_j$$

This is a linear first-order differential equation for the vector of probabilities $\{p_i\}$.

Alternatively, in matrix form:

$$\frac{d\mathbf{p}}{dt} = W\mathbf{p}$$

where W is the **rate matrix** with entries:

- $W_{ij} = \nu_{ji}$ for $i \neq j$,
- $W_{ii} = -\sum_{j \neq i} \nu_{ij}$.

This structure ensures probability conservation: the total probability $\sum_i p_i = 1$ remains constant in time.

The master equation is first order in time and does not have time reversal symmetry so describes an irreversible process. This irreversibility arises from the coarse-graining process that throws away information about underlying microphysics which is described by Newton's equations and which are time reversible. Only by doing so is the entropy allowed to increase which is required by the second law of thermodynamics for an irreversible process. Consequently the increase of entropy is linked to our knowledge about the system rather than anything it is doing internally in a manner that may appear dubious. Can it be possible that macroscopic and reproducible phenomena such as heat flow depend on how we handle information? Perhaps yes since the division between work and heat is somewhat arbitrary. Were we able to track all the particle positions there would be no need to talk about heat energy or heat flow.

11.2 From the Master Equation to the Diffusion Equation

Now consider the situation where the state index i corresponds to a position in space, $x_i = ia$, ie a one-dimensional lattice with lattice spacing a , and transitions only occur between neighboring lattice sites.

We assume:

- Transition rates are symmetric: $\nu_{i,i+1} = \nu_{i,i-1} = \nu$,
- The spacing $a \rightarrow 0$ in the continuum limit.

The master equation becomes a finite-difference equation:

$$\begin{aligned}\frac{dp_i}{dt} &= \sum_j \nu_{ij}(p_j - p_i) \\ \frac{dp_i}{dt} &= \nu(p_{i-1} - p_i) + \nu(p_{i+1} - p_i) \\ \frac{dp_i}{dt} &= \nu(p_{i+1} + p_{i-1} - 2p_i)\end{aligned}$$

We now define a continuous variable $x = ia$, and a probability density $p(x, t)$ such that $p_i(t) \approx p(x, t)$.

Using Taylor expansions:

$$p(x \pm a, t) = p(x, t) \pm a \frac{\partial p}{\partial x} + \frac{a^2}{2} \frac{\partial^2 p}{\partial x^2} + \dots$$

Substituting into the master equation gives:

$$\frac{\partial p}{\partial t} = \nu a^2 \frac{\partial^2 p}{\partial x^2}$$

Defining the diffusion constant $D = \nu a^2$, we obtain the **diffusion equation**:

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}$$

The diffusion equation describes the evolution of the probability density of a particle with diffusion constant (sometimes called diffusivity) given by $D = \nu a^2$. The dimensions of D are $[\text{length}]^2 / [\text{time}]$. Typically, after expansion, we set the lattice spacing a to 1. Additionally, the diffusion equation can describe many non-interacting diffusing particles. In this case, we replace p with ρ , representing the density or concentration of particles, and use the normalization $\int dx \rho = M$, where M is the number of particles.

The diffusion equation, much like the master equation from which it originates, explicitly violates time-reversal symmetry, thus permitting entropy to increase.

The solution of the diffusion equation for an initial condition where the particle is initially localized at the origin (formally, $p(x, 0) = \delta(x)$) is a Gaussian:

$$p(x, t) = (4\pi Dt)^{-1/2} \exp \left[-\frac{x^2}{4Dt} \right]$$

We explicitly see the arrow of time by examining this Gaussian solution at various times t . As t increases, the Gaussian “bell-shaped” curve spreads out. Its width grows according to $\langle x^2 \rangle^{1/2} \sim t^{1/2}$. This is known as “diffusive scaling,” and it implies that, after time t , a particle will typically be found at a distance roughly proportional to $t^{1/2}$ from its starting point. Conversely, exploring a region of size L typically requires a time of order $O(L^2)$.

The evolution of the solution to the 1d diffusion equation in a spatial region $x = [0, 1]$ as a function of times are shown in the movie below. The diffusion constant is $D = 0.01$. The movie corresponds to a particle initialised at $x = 0.5$. One sees how the probability density spreads out over the range as time increases. This can be used to model the diffusion of particles down a concentration gradient as you will see in the next part of the course.

[Movies/diffusion_evolution.mp4](#)

```
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.animation as animation
import os

# Parameters
L = 1.0          # Length of the domain
T = 1.0          # Total time
nx = 400         # Number of spatial points
```

```

nt = 2000      # Number of time steps
D = 0.1        # Diffusion coefficient

# Discretization
dx = L / (nx - 1)
dt = T / nt

# Stability condition auto-adjust
if D * dt / dx**2 > 0.5:
    print("Adjusting dt and nt to satisfy stability condition...")
    dt = 0.4 * dx**2 / D
    nt = int(T / dt)
    dt = T / nt # Recalculate dt exactly

nt = int(nt / 50) # Artificial slowdown for animation

print(f"Using dt = {dt:.4e}, nt = {nt}")

# Initialize x and u
x = np.linspace(0, L, nx)

# Initial condition: smooth narrow Gaussian
sigma = 0.01
u = np.exp(-(x - L/2)**2 / (2 * sigma**2))

# Normalize initial condition
u /= np.sum(u) * dx

# Setup figure
plt.rcParams.update({
    "text.usetex": True,
    "font.family": "serif"
})

fig, ax = plt.subplots()
line, = ax.plot(x, u)

# Compute peak height for setting y-axis
peak_height = 1 / (np.sqrt(2 * np.pi) * sigma)
ax.set_ylim(0, peak_height * 1.05)

ax.set_xlabel(r'$x$', fontsize=20)
ax.set_ylabel(r'$p(x,t)$', fontsize=20)

ax.tick_params(axis='both', which='major', labelsize=14)

# Tiny time counter text
time_text = ax.text(0.85, 0.05, '', transform=ax.transAxes, fontsize=10, verticalalignment='bottom')

# Function to update the plot
def update(frame):
    global u

```

```

unew = np.copy(u)
unew[1:-1] = u[1:-1] + D * dt / dx**2 * (u[2:] - 2*u[1:-1] + u[:-2])
u = unew

# Normalize at every step
u /= np.sum(u) * dx

line.set_ydata(u)
current_time = frame * dt
time_text.set_text(r'$t=%4f$' % current_time)
return line, time_text

ani = animation.FuncAnimation(fig, update, frames=nt, interval=100, blit=True)

# Save the animation as a movie
Writer = animation.writers['ffmpeg']
writer = Writer(fps=15, metadata=dict(artist='Me'), bitrate=1800)
ani.save("../Movies/diffusion_evolution.mp4", writer=writer)

plt.show()

```

11.3 Consequences of time reversal symmetry

As we have seen by introducing a type of coarse graining, the master equation violates time reversal symmetry of the underlying Newtonian dynamics. Remarkably however, the fact that the underlying microphysics is actually time reversal symmetric has several deep consequences which survive the coarse graining procedure. These results are some of the cornerstones of nonequilibrium thermodynamics.

11.3.1 Detailed balance

Recall from year 2 Statistical Mechanics that for a system in equilibrium, the principle of equal a-priori probabilities of microstates holds. Therefore

$$\nu_{ij} p_i^{eq} = \nu_{ji} p_j^{eq}$$

Hence, on average, the actual rate of quantum jumps from i to j (the left-hand side) is the same as from j to i . This is a stronger statement than the master equation, which asserts only that there is overall balance between the rate of jumping into and out of state i in equilibrium. The result above is known as the *principle of detailed balance*.

This principle is powerful because it applies not only to individual states but also to any grouping of states.

Exercise: Show that for two groups of states, A and B , the overall rate of transitions from group A to group B is balanced, in equilibrium, by those from B to A :

$$\nu_{AB} p_A^{eq} = \nu_{BA} p_B^{eq}$$

Hence, detailed balance arguments can be extended to subsystems within a large isolated system, and even to systems that are not isolated. However, in such cases, the principle is far from obvious, because once states are grouped together:

$$\nu_{AB} \neq \nu_{BA}, \quad p_A \neq p_B$$

(This can be easily demonstrated, for example, by considering two groups that contain different numbers of states with similar energies.) Nonetheless, the detailed balance relation holds, in equilibrium, in the general form above.

11.3.2 Computer simulation

In computer simulation, good results will be obtained if one accurately follows the microscopic equations of motion. This is the *molecular dynamics (MD)* method which we now outline.

Molecular dynamics

Molecular Dynamics (MD) involves a system of classical particles interacting through specified interparticle forces. The motion of these particles is determined by numerically integrating Newton's equations of motion. In MD simulations, averages of state variables are obtained as time averages over trajectories in phase space. Typically, the forces acting between particles are conservative, ensuring that the total energy E remains constant. This conservation implies that the motion is restricted to a $(2dN - 1)$ -dimensional surface in phase space, denoted by $\Gamma(E)$.

A central aspect of MD is the averaging of observables. For a given observable A , its average is computed as the time average along the trajectory. Mathematically, this is expressed as:

$$\langle A(\{\mathbf{p}_i\}, \{\mathbf{r}_i\}) \rangle = \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} dt A(\{\mathbf{p}_i(t)\}, \{\mathbf{r}_i(t)\})$$

This formula represents the integral of the observable over a time interval τ , normalized by the length of that interval.

The practical steps of an MD simulation start with generating an initial random configuration of particle positions $\{\mathbf{r}_i\}$ and momenta $\{\mathbf{p}_i\}$. The system's equations of motion are then iteratively solved using a suitable algorithm to allow it to reach equilibrium. After equilibration, a production run is performed over many time steps to collect meaningful data. Finally, relevant averages, such as pressure or kinetic energy, are calculated from the collected data.

Monte Carlo

However, to obtain the equilibrium properties of the system, it may be much faster to use a dynamics which is nothing like the actual equations of motion.

At first sight, this looks very dangerous; however, if one can prove that in the required equilibrium distribution, the artificial dynamics obey the principle of detailed balance, then it is (almost) guaranteed that the steady state found by simulation is the true equilibrium state.

The best known example is the *Monte Carlo method*, in which the dynamical algorithm consists of random jumps. The jump rates ν_{AB} for all pairs of states (A, B) take the form:

$$\nu_{AB} = \nu_0 \quad \text{if } E_B \leq E_A$$

$$\nu_{AB} = \nu_0 e^{-\beta(E_B - E_A)} \quad \text{if } E_B \geq E_A$$

where ν_0 is a constant.

Exercise: Show that this gives the canonical distribution in steady state.

🔥 Solution

To show that this jump rate rule gives the canonical distribution in steady state, we assume the system reaches a steady-state probability distribution P_A for state A and apply the condition of detailed balance.

Detailed balance requires:

$$P_A \nu_{AB} = P_B \nu_{BA}$$

Assume the canonical distribution:

$$P_A = \frac{1}{Z} e^{-\beta E_A}, \quad P_B = \frac{1}{Z} e^{-\beta E_B}$$

Now consider two cases for ν_{AB} and ν_{BA} :

Case 1: $E_B \leq E_A$

Then:

$$\nu_{AB} = \nu_0, \quad \nu_{BA} = \nu_0 e^{-\beta(E_A - E_B)}$$

Substitute into the detailed balance condition:

$$P_A \nu_0 = P_B \nu_0 e^{-\beta(E_A - E_B)}$$

Cancel ν_0 :

$$P_A = P_B e^{-\beta(E_A - E_B)}$$

Use canonical form:

$$\frac{1}{Z} e^{-\beta E_A} = \frac{1}{Z} e^{-\beta E_B} e^{-\beta(E_A - E_B)} = \frac{1}{Z} e^{-\beta E_A}$$

Verified.

Case 2: $E_B > E_A$

Then:

$$\nu_{AB} = \nu_0 e^{-\beta(E_B - E_A)}, \quad \nu_{BA} = \nu_0$$

Substitute:

$$P_A \nu_0 e^{-\beta(E_B - E_A)} = P_B \nu_0$$

Cancel ν_0 :

$$P_A e^{-\beta(E_B - E_A)} = P_B$$

Again using canonical form:

$$\frac{1}{Z} e^{-\beta E_A} e^{-\beta(E_B - E_A)} = \frac{1}{Z} e^{-\beta E_B} = P_B$$

Verified.

Hence, in both cases the detailed balance condition is satisfied with the canonical distribution, and the steady state is indeed the canonical distribution.

Chapter 12

The Langevin Approach

12.1 The Random Walk and the Langevin equation

The concept of a random walk and its continuum limit – diffusion – introduced in the previous chapter, expresses the time evolution of the probability distribution $p(x, t)$ for a particle's position x by the diffusion equation:

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2},$$

which is a standard example of a so called *Fokker-Planck* equation, which is second-order in space and first-order in time.

In contrast, the *Langevin equation* provides a stochastic differential equation for the particle's trajectory $x(t)$. To understand it, consider the random hopping motion of a particle on a 1d lattice over a small time increment Δt :

$$x(t + \Delta t) = x(t) + \Delta x(t)$$

Here, $\Delta x(t)$ is a random displacement. If the lattice spacing is a , we define the step statistics as:

$$\Delta x(t) = \begin{cases} +a & \text{with probability } \nu \Delta t \\ -a & \text{with probability } \nu \Delta t \\ 0 & \text{with probability } 1 - 2\nu \Delta t \end{cases}$$

This defines a discrete-time, discrete-space random walk. The average and variance of the step are easily derived using binomial statistics:

- Mean: $\langle \Delta x \rangle = 0$
- Variance: $\langle (\Delta x)^2 \rangle = 2a^2 \nu \Delta t = 2D \Delta t$

The steps $\Delta x(t)$ are uncorrelated across time.

To take the continuum limit, we let both $a \rightarrow 0$ and $\Delta t \rightarrow 0$ in such a way that:

$$a \propto \sqrt{\Delta t}$$

In this limit, we obtain the **Langevin equation**:

$$\dot{x}(t) = \eta(t)$$

where $\eta(t) \equiv \frac{\Delta x(t)}{\Delta t}$ is a stochastic noise satisfying:

$$\langle \eta(t) \rangle = 0$$

$$\langle \eta(t)\eta(t') \rangle = \Gamma\delta(t-t')$$

This $\eta(t)$ is known as white noise — it has zero mean and is uncorrelated at different times. Γ measures its amplitude.

The Langevin equation tells us that the velocity $\dot{x}(t)$ is purely driven by noise. We can formally integrate it:

$$x(t) - x_0 = \int_0^t \eta(t') dt'$$

Taking ensemble averages:

- Mean displacement:

$$\langle x(t) - x_0 \rangle = 0$$

- Mean square displacement:

$$\langle [x(t) - x_0]^2 \rangle = \int_0^t \int_0^t \langle \eta(t')\eta(t'') \rangle dt' dt'' = \Gamma \int_0^t dt' = \Gamma t$$

Comparing this with the diffusion equation result, we identify:

$$\Gamma = 2D$$

Hence, the Langevin description yields the same physical behavior — not just the mean-square displacement but also the full probability distribution $p(x, t)$ — as the diffusion (Fokker-Planck) equation. This equivalence arises from the fact that the integral of many small, independent random steps leads to a Gaussian distribution, in agreement with the solution of the diffusion equation.

For more details, see: *Stochastic Processes in Physics and Chemistry* by N.G. van Kampen (North Holland, 1981).

12.2 Brownian Motion

Let us now examine Brownian motion, originally observed as the erratic motion of colloidal particles suspended in a fluid. These particles undergo constant collisions with surrounding (smaller) fluid molecules, which results in seemingly random movement.

From a coarse-grained perspective — where we do not track each individual collision — this appears as motion under random forces. This statistical treatment introduces irreversibility at the macroscopic level, even though the underlying molecular dynamics are reversible.

The Langevin equation provides a way to model this behavior. For a particle of mass m in one dimension, Langevin proposed the equation:

$$m\ddot{x} = -\gamma\dot{x} + f(t)$$

Here:

- $-\gamma\dot{x}$ is a frictional damping force, where γ is the damping coefficient.
- $f(t)$ is a random force due to molecular collisions.

Often, the mobility is defined as $\mu = 1/\gamma$ — note that this is unrelated to chemical potential.

12.2.1 Noise Properties

In principle the random forces are correlated in time since the molecular collisions which cause them are correlated and have some definite duration.

Let us assume that there is some correlation time t_c over which $\langle f(t_1)f(t_2) \rangle = g(t_1 - t_2)$ decays rapidly as shown in the sketch below:

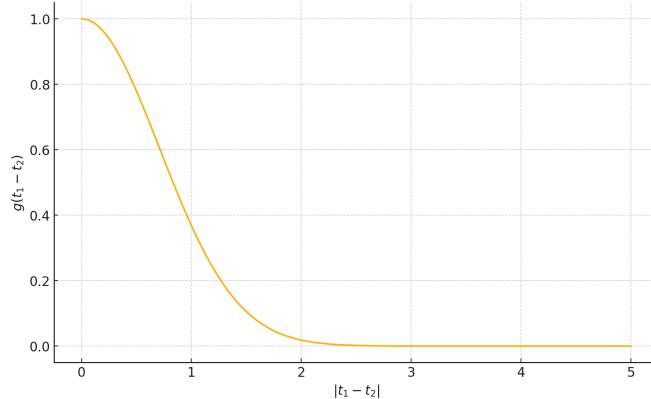


Figure 12.1: Sketch of $g(t_1 - t_2)$ against $|t_1 - t_2|$

Then as long as we consider timescales $\gg t_c$ we can safely replace $g(t_1 - t_2)$ by a delta function. Thus we can make the approximation of white noise

$$\langle f(t) \rangle = 0$$

$$\langle f(t_1)f(t_2) \rangle = \Gamma \delta(t_1 - t_2)$$

12.2.2 Solving the Langevin Equation (velocity)

Let's set $m = 1$ for simplicity and solve the equation:

$$\dot{v} + \gamma v = f(t)$$

We apply an integrating factor:

$$\frac{d}{dt} [ve^{\gamma t}] = e^{\gamma t} f(t)$$

Integrating both sides:

$$v(t) = v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-t')} f(t') dt'$$

Taking the average:

$$\langle v(t) \rangle = v_0 e^{-\gamma t}$$

Thus:

- At short times: ($\gamma t \ll 1$): $\langle v \rangle \approx v_0$ ie. friction is negligible.
- At long times: ($\gamma t \gg 1$): $\langle v \rangle \rightarrow 0$ ie. the system loses memory of the initial velocity.

12.2.3 Mean-square velocity

We now compute:

$$\langle v(t)^2 \rangle = v_0^2 e^{-2\gamma t} + \Gamma \int_0^t e^{-2\gamma(t-t')} dt' = v_0^2 e^{-2\gamma t} + \frac{\Gamma}{2\gamma} (1 - e^{-2\gamma t})$$

Implying that at

- Short times: $\langle v^2 \rangle \approx v_0^2$
- Long times: $\langle v^2 \rangle \rightarrow \Gamma/(2\gamma)$

At equilibrium, the equipartition theorem gives:

$$\frac{1}{2}m\langle v^2 \rangle = \frac{1}{2}k_B T$$

Using this to identify Γ :

$$\Gamma = 2\gamma k_B T$$

This important result relates the noise strength to the damping and temperature — they have the same microscopic origin (molecular collisions).

12.2.4 Mean-square displacement

We now integrate $v(t)$ again to get position $x(t)$ (with $m = 1$):

Using the result above and substituting $\Gamma = 2\gamma k_B T$, we find:

$$\langle [x(t) - x_0]^2 \rangle = \frac{(v_0^2 - k_B T)}{\gamma^2} (1 - e^{-\gamma t})^2 + \frac{2k_B T}{\gamma} \left[t - \frac{1 - e^{-\gamma t}}{\gamma} \right]$$

Limiting behaviours: