2 Some necessary background

2.1 THERMODYNAMICS AND STATISTICAL MECHANICS: A QUICK REMINDER

2.1.1 Basic notions

In this chapter we shall review some of the basic features of thermodynamics and statistical mechanics which will be used later in this book when devising simulation methods and interpreting results. Many good books on this subject exist and we shall not attempt to present a complete treatment. This chapter is hence *not* intended to *replace* any textbook for this important field of physics but rather to 'refresh' the reader's knowledge and to draw attention to notions in thermodynamics and statistical mechanics which will henceforth be assumed to be known throughout this book.

2.1.1.1 Partition function

Equilibrium statistical mechanics is based upon the idea of a partition function which contains all of the essential information about the system under consideration. The general form for the partition function for a classical system is

$$Z = \sum_{\text{all states}} e^{-\mathcal{H}/k_{\mathrm{B}}T}, \tag{2.1}$$

where $\mathcal{H}$ is the Hamiltonian for the system, T is the temperature, and k_{B} is the Boltzmann constant. The sum in Eqn. (2.1) is over all possible states of the system and thus depends upon the size of the system and the number of degrees of freedom for each particle. For systems consisting of only a few interacting particles the partition function can be written down exactly with the consequence that the properties of the system can be calculated in closed form. In a few other cases the interactions between particles are so simple that evaluating the partition function is possible. Of course this means that we have only adopted classical Hamiltonians when describing the potential energy of a system rather than using a quantum mechanical operator.

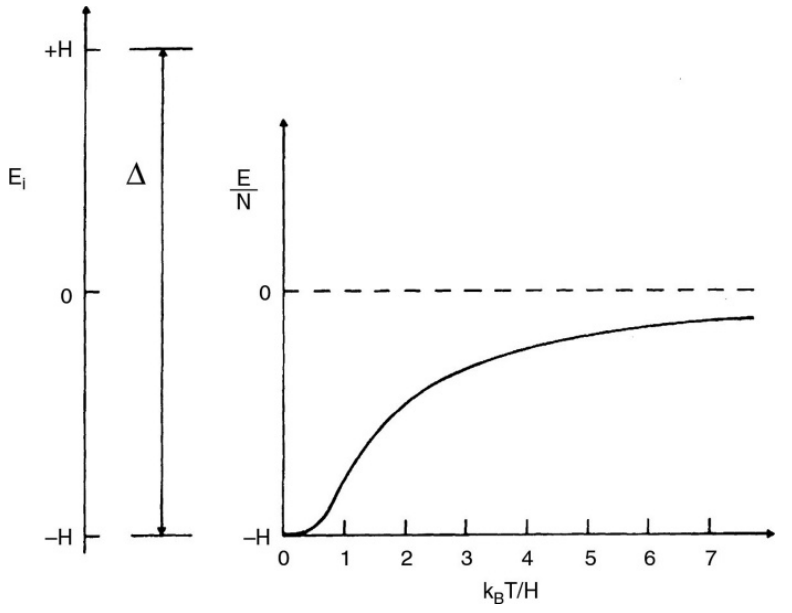


Fig. 2.1 (left) Energy levels for the two-level system in Eqn. (2.2); (right) internal energy for a two-level system as a function of temperature.

Example

Let us consider a system with N particles each of which has only two states, e.g. a non-interacting Ising model in an external magnetic field H, and which has the Hamiltonian

$$\mathcal{H} = -H \sum_i \sigma_i, \tag{2.2}$$

where $\sigma_i = \pm 1$. The partition function for this system is simply

$$Z = \left(e^{-H/k_B T} + e^{+H/k_B T} \right)^N, \tag{2.3}$$

where for a single spin the sum in Eqn. (2.1) is only over two states. The energies of the states and the resultant temperature dependence of the internal energy appropriate to this situation are pictured in Fig. 2.1.

Problem 2.1 Work out the average magnetization per spin, using Eqn. (2.3), for a system of N non-interacting Ising spins in an external magnetic field. [Solution $M = -(1/N)\partial F/\partial H$, $F = -k_B T \ln Z \Rightarrow M = \tanh(H/k_B T)$]

There are also a few examples where it is possible to extract exact results for very large systems of interacting particles, but in general the partition function cannot be evaluated exactly. Even enumerating the terms in the partition function on a computer can be a daunting task. Even if we have only 10 000 interacting particles, a very small fraction of Avogadro's number, with only two possible states per particle, the partition function would contain $2^{10\,000}$

terms! The probability of any particular state of the system is also determined by the partition function. Thus, the probability that the system is in state μ is given by

$$P_\mu = e^{-\mathcal{H}(\mu)/k_B T}/Z, \tag{2.4}$$

where $\mathcal{H}(\mu)$ is the Hamiltonian when the system is in the μth state. As we shall show in succeeding chapters, the Monte Carlo method is an excellent technique for estimating probabilities, and we can take advantage of this property in evaluating the results.

2.1.1.2 Free energy, internal energy, and entropy

It is possible to make a direct connection between the partition function and thermodynamic quantities and we shall now briefly review these relationships. The free energy of a system can be determined from the partition function (Callen, 1985) from

$$F = -k_B T \ln Z \tag{2.5}$$

and all other thermodynamic quantities can be calculated by appropriate differentiation of Eqn. (2.5). This relation then provides the connection between statistical mechanics and thermodynamics. The internal energy of a system can be obtained from the free energy via

$$U = -T^2 \partial(F/T)/\partial T. \tag{2.6}$$

By the use of a partial derivative we imply here that F will depend upon other variables as well, e.g. the magnetic field H in the above example, which are held constant in Eqn. (2.6). This also means that if the internal energy of a system can be measured, the free energy can be extracted by appropriate integration, assuming, of course, that the free energy is known at some reference temperature. We shall see that this fact is important for simulations which do not yield the free energy directly but produce instead values for the internal energy. Free energy differences may then be estimated by integration, i.e. from $\Delta(F/T) = \int d(1/T)U$.

Using Eqn. (2.6) one can easily determine the temperature dependence of the internal energy for the non-interacting Ising model, and this is also shown in Fig. 2.1. Another important quantity, the entropy, measures the amount of disorder in the system. The entropy is defined in statistical mechanics by

$$S = -k_B \ln P, \tag{2.7}$$

where P is the probability of occurrence of a (thermodynamic) microstate. The entropy can be determined from the free energy from

$$S = -(\partial F/\partial T)_{V,N}. \tag{2.8}$$

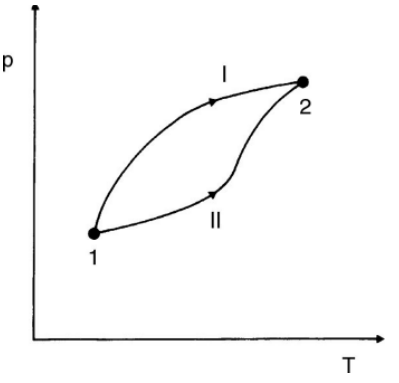


Fig. 2.2 Schematic view of different paths between two different points in thermodynamic $p-T$ space.

2.1.1.3 Thermodynamic potentials and corresponding ensembles

The internal energy is expressed as a function of the extensive variables, S, V, N, etc. There are situations when it is appropriate to replace some of these variables by their conjugate intensive variables, and for this purpose additional thermodynamic potentials can be defined by suitable Legendre transforms of the internal energy; in terms of liquid–gas variables such relations are given by:

$$F = U - TS, \tag{2.9a}$$

$$H = U + pV, \tag{2.9b}$$

$$G = U - TS + pV, \tag{2.9c}$$

where F is the Helmholtz free energy, H is the enthalpy, and G is the Gibbs free energy. Similar expressions can be derived using other thermodynamic variables, e.g. magnetic variables. The free energy is important since it is a minimum in equilibrium when T and V are held constant, while G is a minimum when T and p are held fixed. Moreover, the difference in free energy between any two states does not depend on the path between the states. Thus, in Fig. 2.2 we consider two points in the $p-T$ plane. Two different paths which connect points 1 and 2 are shown; the difference in free energy between these two points is identical for both paths, i.e.

$$F_2 - F_1 = \int_{\text{path I}} dF = \int_{\text{path II}} dF. \tag{2.10}$$

The multidimensional space in which each point specifies the complete microstate (specified by the degrees of freedom of all the particles) of a system is termed 'phase space'. Averages over phase space may be constructed by considering a large number of identical systems which are held at the same fixed conditions. These are called 'ensembles'. Different ensembles are relevant for different constraints. If the temperature is held fixed, the set of systems is said to belong to the 'canonical ensemble' and there will be some distribution of energies among the different systems. If instead the energy is fixed, the ensemble is termed the 'microcanonical' ensemble. In the first two cases the

number of particles is held constant; if the number of particles is allowed to fluctuate the ensemble is the 'grand canonical' ensemble.

Systems are often held at fixed values of intensive variables, such as temperature, pressure, etc. The conjugate extensive variables, energy, volume, etc. will fluctuate with time; indeed these fluctuations will actually be observed during Monte Carlo simulations.

It is important to recall that Legendre transformations from one thermodynamic potential to another one (e.g. from the microcanonical ensemble, where $U(S, V)$ is a function of its 'natural variables' S and V, to the canonical ensemble where $F(T, V)$ is considered as a function of the 'natural variables' T and V) are only fully equivalent to each other in the thermodynamic limit, $N \to \infty$. For finite N it is still true that, in thermal equilibrium, $F(T, V, N)$ is a minimum at fixed T and V, and $U(S, V, N)$ is minimized at fixed S and V, but Eqn. (2.9) no longer holds for finite N since finite size effects in different ensembles are no longer equivalent. This is particularly important when considering phase transitions (see Section 2.1.2). However, on the level of partition functions and probability distributions of states there are exact relations between different ensembles. For example, the partition function $Y(\mu, V, T)$ of the grand-canonical ensemble (where the chemical potential μ is a 'natural variable' rather than N) is related to the canonical partition function $Z(N, V, T)$ by

$$Y(\mu, V, T) = \sum_{N=0}^{\infty} \exp(\mu N / k_B T) Z(N, V, T), \tag{2.11a}$$

and the probability to find the system in a particular state $\vec{X}$ with N particles is ($\vec{X}$ stands for the degrees of freedom of these particles)

$$P_{\mu VT}(\vec{X}, N) = (1/Y) \exp(\mu N / k_B T) \exp[-U(\vec{X}) / k_B T]. \tag{2.11b}$$

In the canonical ensemble, where N does not fluctuate, we have instead

$$P_{NVT}(\vec{X}) = (1/Z) \exp[-U(\vec{X}) / k_B T]. \tag{2.11c}$$

In the limit of macroscopic systems ($V \to \infty$) the distribution $P_{\mu VT}(\vec{X}, N)$ essentially becomes a δ-function in N, sharply peaked at $\overline{N} = \sum_N P_{\mu VT}(\vec{X}, N) N$. Then, averages in the canonical and grand-canonical ensembles become strictly equivalent, and the potentials are related via Legendre transformations ($\Omega = -k_B T \ln Y = F - \mu N$). For 'small' systems, such as those studied in Monte Carlo simulations, use of Legendre transformations is only useful if finite size effects are negligible.

In the context of the study of phase transitions and phase coexistence, it is sometimes advantageous to complement studies of a system in the canonical (NVT) ensemble that has been emphasized so far in this book, by studying the system in the microcanonical (NVE) ensemble. If we include the kinetic energy $K(\{\vec{p}_i\})$ of the particles in the discussion, the Hamiltonian becomes $\mathcal{H}(\{\vec{x}_i\}, \{\vec{p}_i\}, V) = K(\{\vec{p}_i\}) + U(\{\vec{x}_i\}, V)$. Including the kinetic energy is necessary, of course, when discussing molecular dynamics simulation methods, since solving Newton's equation of motion conserves the total energy $E = \mathcal{H}(\{\vec{x}_i\},$

$\{\vec{p}_i\}, V$). Here we denote the coordinates of the particles in the d-dimensional volume by $\{\vec{x}_i, i = 1, \ldots, N\}$, and $\{\vec{p}_i\}$ are the momenta.

The basic thermodynamic potential then is the entropy

$$S = S(E, V, N) = k_B \ln Z_{MC} = k_B \ln \Omega(E, V, N), \tag{2.12}$$

where the microcanonical partition function Z_{MC} is nothing but the phase space volume $\Omega(E, V, N)$ defined by (Becker, 1967; Dunkel and Hilbert, 2006):

$$\Omega(E, V, N) = \frac{1}{N! h^{dN}} \int dx \int dp\, \Theta[E - \mathcal{H}(\{\vec{x}_i\}, \{\vec{p}_i\}, V)], \tag{2.13}$$

where h is Planck's constant, $\Theta(z)$ is Heaviside's step function $\{\Theta(z < 0) = 0, \Theta(z \geq 0) = 1\}$, and the integrals $\int dx$, $\int dp$ stand symbolically for N d-dimensional integrals over the components of coordinates and momenta of the particles. Note that due to the high dimensionality of phase space for large N almost all the weight for this integration comes from the surface of the integration volume, and hence often an expression, equivalent for $N \to \infty$,

$$Z_{MC} = \varepsilon_0 \frac{\partial \Omega}{\partial E} = \frac{1}{N! h^{dN}} \varepsilon_0 \int dx \int dp\, \delta[E - \mathcal{H}(\{\vec{x}_i\}, \{\vec{p}_i\}, V)], \tag{2.14}$$

is quoted in textbooks (wher ε_0 is the thickness of a thin energy shell around the phase space surface defined by $\mathcal{H}(\{\vec{x}_i\}, \{\vec{p}_i\}, V) = E$). However, Eqn. (2.13) is preferable since it also holds for small N, and the unspecified parameter ε_0 is not needed. $\partial\Omega/\partial E$ $\partial\Omega$ is proportional to the microcanonical energy density of states.

If $S(E, V, N)$ is given, other thermodynamic variables are found as usual from derivatives

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V}, \qquad \frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{N,T} \tag{2.15}$$

In the thermodynamic limit, one can show that $S(E, V, N)$ is a convex function of its variables, e.g. the energy E. This fact is of particular interest if we consider a system undergoing a first order phase transition between two phases I and II. As described in Section 2.1 in the canonical ensemble we then find that both the total internal energy U and the entropy S undergo jumps at the transition temperature T_t, from U_I to U_{II} and from S_I to S_{II}, respectively. (Remember that for classical systems, momenta cancel out from the canonic partition function, so kinetic energy does not need to be considered.) Of course, for $N \to \infty$ all statistical ensembles are equivalent, related by Legendre transformations, and hence in the microcanonical ensemble the first order transition shows up as a linear variation of $S(U)$ from S_I (U_I) to S_{II} (U_{II}), and $1/T$ (Eqn. (2.15)) is then simply a constant inbetween U_I and U_{II}. This physical meaning of this linear variation is that the system passes through a two-phase coexistence region for which the simple lever rule holds.

In a finite system, the linear variation in $S(U)$ typically is replaced by a concave intruder, and the constant piece in the $1/T$ vs. U curve is replaced by a loop. However, these observations should not be interpreted in terms of

concepts such as Landau potentials, van der Waals loops, and the like; rather for large but finite N these phenomena can be attributed to interfacial effects on phase coexistence (although in the thermodynamic potential, the entropy in the present case, these effects are small, down by a surface to volume ratio, they often can be rather easily recorded). In any case, these 'intruders' and 'loops' are useful indicators for the presence of a first order phase transition and can be used to characterize them precisely.

Problem 2.2 Consider a two-level system composed of N non-interacting particles where the groundstate of each particle is doubly degenerate and separated from the upper level by an energy ΔE. What is the partition function for this system? What is the entropy as a function of temperature?

2.1.1.4 Fluctuations

Equations (2.4) and (2.5) imply that the probability that a given 'microstate' μ occurs is $P_\mu = \exp\{[F - \mathcal{H}(\mu)]/k_\mathrm{B}T\} = \exp\{-S/k_\mathrm{B}\}$. Since the number of different microstates is so huge, we are not only interested in probabilities of individual microstates but also in probabilities of macroscopic variables, such as the internal energy U. We first form the moments (where $\beta \equiv k_\mathrm{B}T$; the average energy is denoted $\overline{U}$ and U is a fluctuating quantity),

$$\begin{aligned} \overline{U}(\beta) = \langle \mathcal{H}(\mu) \rangle \equiv \sum_\mu P_\mu \mathcal{H}(\mu) = \sum_\mu \mathcal{H}(\mu) e^{-\beta \mathcal{H}(\mu)} \Big/ \sum_\mu e^{-\beta \mathcal{H}(\mu)}, \\ \langle \mathcal{H}^2 \rangle = \sum_\mu \mathcal{H}^2 e^{-\beta \mathcal{H}(\mu)} \Big/ \sum_\mu e^{-\beta \mathcal{H}(\mu)}, \end{aligned} \tag{2.16}$$

and note the relation $-(\partial U(\beta)/\partial\beta)_V = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2$. Since $(\partial U/\partial T)_V = C_V$, the specific heat thus yields a fluctuation relation

$$k_\mathrm{B}T^2 C_V = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = \langle (\Delta U)^2 \rangle_{NVT}, \quad \Delta U \equiv \mathcal{H} - \langle \mathcal{H} \rangle. \tag{2.17}$$

Now for a macroscopic system ($N \gg 1$) away from a critical point, $U \propto N$ and the energy and specific heat are extensive quantities. However, since both $\langle \mathcal{H}^2 \rangle$ and $\langle \mathcal{H} \rangle^2$ are clearly proportional to N^2, we see that the relative fluctuation of the energy is very small, of order $1/N$. While in real experiments (where often $N \approx 10^{22}$) such fluctuations may be too small to be detectable, in simulations these thermal fluctuations are readily observable, and relations such as Eqn. (2.17) are useful for the actual estimation of the specific heat from energy fluctuations. Similar fluctuation relations exist for many other quantities, e.g. the isothermal susceptibility $\chi = (\partial \langle M \rangle / \partial H)_T$ is related to fluctuations of the magnetization $M = \sum_i \sigma_i$, as

$$k_\mathrm{B}T\chi = \langle M^2 \rangle - \langle M \rangle^2 = \sum_{ij} (\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle). \tag{2.18}$$

Writing the Hamiltonian of a system in the presence of a magnetic field H as $\mathcal{H} = \mathcal{H}_0 - HM$, we can easily derive Eqn. (2.18) from $\langle M \rangle =$

$\sum_\mu M \exp[-\beta\mathcal{H}(\mu)]/\sum_\mu \exp[-\beta\mathcal{H}(\mu)]$ in a similar fashion as above. The relative fluctuation of the magnetization is also small, of order $1/N$.

It is not only of interest to consider for quantities such as the energy or magnetization the lowest order moments but to discuss the full probability distribution $P(U)$ or $P(M)$, respectively. For a system in a pure phase the probability is given by a simple Gaussian distribution

$$P(U) = (2\pi k_B C_V T^2)^{-1/2} \exp[-(\Delta U)^2/2k_B T^2 C_V] \tag{2.19}$$

while the distribution of the magnetization for the paramagnetic system becomes

$$P(M) = (2\pi k_B T\chi)^{-1/2} \exp[-(M - \langle M\rangle)^2/2k_B T\chi] \tag{2.20}$$

It is straightforward to verify that Eqns. (2.19) and (2.20) are fully consistent with the fluctuation relations (2.17) and (2.18). Since Gaussian distributions are completely specified by the first two moments, higher moments $\langle H^k\rangle$, $\langle M^k\rangle$, which could be obtained analogously to Eqn. (2.16), are not required. Note that on the scale of $\overline{U}/N$ and $\langle M\rangle/N$ the distributions $P(U)$, $P(M)$ are extremely narrow, and ultimately tend to δ-functions in the thermodynamic limit. Thus these fluctuations are usually neglected altogether when dealing with relations between thermodynamic variables.

An important consideration is that the thermodynamic state variables do not depend on the ensemble chosen (in pure phases) while the fluctuations do. Therefore, one obtains the same average internal energy $\overline{U}(N, V, T)$ in the canonical ensemble as in the Np_T ensemble while the specific heats and the energy fluctuations differ (see Landau and Lifshitz, 1980):

$$\langle(\Delta U)^2\rangle_{Np_T} = k_B T^2 C_V - \left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]^2 k_B T \left(\frac{\partial V}{\partial p}\right)_T. \tag{2.21}$$

It is also interesting to consider fluctuations of several thermodynamic variables together. Then one can ask whether these quantities are correlated, or whether the fluctuations of these quantities are independent of each other. Consider the NVT ensemble where entropy S and the pressure p (an intensive variable) are the (fluctuating) conjugate variables $\{p = -(\partial F/\partial V)_{NT}$, $S = -(\partial F/\partial T)_{NV}\}$. What are the fluctuations of S and p, and are they correlated? The answer to these questions is given by

$$\langle(\Delta S)^2\rangle_{NVT} = k_B C_p, \tag{2.22a}$$

$$\langle(\Delta p)^2\rangle_{NVT} = -k_B T(\partial p/\partial V)_S, \tag{2.22b}$$

$$\langle(\Delta S)(\Delta p)\rangle_{NVT} = 0. \tag{2.22c}$$

One can also see here an illustration of the general principle that fluctuations of extensive variables (like S) scale with the volume, while fluctuations of intensive variables (like p) scale with the inverse volume.

2.1.2 Phase transitions

The emphasis in the standard texts on statistical mechanics clearly is on those problems that can be dealt with analytically, e.g. ideal classical and quantum gases, dilute solutions, etc. The main utility of Monte Carlo methods is for problems which evade exact solution such as phase transitions, calculations of phase diagrams, etc. For this reason we shall emphasize this topic here. The study of phase transitions has long been a topic of great interest in a variety of related scientific disciplines and plays a central role in research in many fields of physics. Although very simple approaches, such as mean field theory, provide a very simple, intuitive picture of phase transitions, they generally fail to provide a quantitative framework for explaining the wide variety of phenomena which occur under a range of different conditions and often do not really capture the conceptual features of the important processes which occur at a phase transition. The last half century has seen the development of a mature framework for the understanding and classification of phase transitions using a combination of (rare) exact solutions as well as theoretical and numerical approaches.

We draw the reader's attention to the existence of zero temperature quantum phase transitions (Sachdev, 1999). These are driven by control parameters that modify the quantum fluctuations and can be studied using quantum Monte Carlo methods that will be described in Chapter 8. The discussion in this chapter, however, will be limited to classical statistical mechanics.

2.1.2.1 Order parameter

The distinguishing feature of most phase transitions is the appearance of a non-zero value of an 'order parameter', i.e. of some property of the system which is non-zero in the ordered phase but identically zero in the disordered phase. The order parameter is defined differently in different kinds of physical systems. In a ferromagnet it is simply the spontaneous magnetization. In a liquid–gas system it will be the difference in the density between the liquid and gas phases at the transition; for liquid crystals the degree of orientational order is telling. An order parameter may be a scalar quantity or may be a multicomponent (or even complex) quantity. Depending on the physical system, an order parameter may be measured by a variety of experimental methods such as neutron scattering, where Bragg peaks of super-structures in antiferromagnets allow the estimation of the order parameter from the integrated intensity, oscillating magnetometer measurement directly determines the spontaneous magnetization of a ferromagnet, while NMR is suitable for the measurement of local orientational order.

2.1.2.2 Correlation function

Even if a system is not ordered, there will, in general, be microscopic regions in the material in which the characteristics of the material are correlated.

Correlations are generally measured through the determination of a two-point correlation function

$$\Gamma(r) = \langle \rho(0)\rho(r) \rangle \,, \tag{2.23}$$

where r is the spatial distance and ρ is the quantity whose correlation is being measured. (The behavior of this correlation function will be discussed shortly.) It is also possible to consider correlations that are both space-dependent and time-dependent, but at the moment we only consider equal time correlations that are time-independent. As a function of distance they will decay (although not always monotonically), and if the correlation for the appropriate quantity decays to zero as the distance goes to infinity, then the order parameter is zero.

2.1.2.3 First order vs. second order

These remarks will concentrate on systems which are in thermal equilibrium and which undergo a phase transition between a disordered state and one which shows order which can be described by an appropriately defined order parameter. If the first derivatives of the free energy are discontinuous at the transition temperature T_c, the transition is termed first order. The magnitude of the discontinuity is unimportant in terms of the classification of the phase transition, but there are diverse systems with either very large or rather small 'jumps'. For second order phase transitions first derivatives are continuous; transitions at some temperature T_c and 'field' H are characterized by singularities in the second derivatives of the free energy, and properties of rather disparate systems can be related by considering not the absolute temperature, but rather the reduced distance from the transition $\varepsilon = |1 - T/T_c|$. (Note that in the 1960s and early 1970s the symbol ε was used to denote the reduced distance from the critical point. As renormalization group theory came on the scene, and in particular ε-expansion techniques became popular, the notation changed to use the symbol t instead. In this book, however, we shall often use the symbol t to stand for time, so to avoid ambiguity we have returned to the original notation.) In Fig. 2.3 we show characteristic behavior for both kinds of phase transitions. At a first order phase transition the free energy curves for ordered and disordered states cross with a finite difference in slope and both stable and metastable states exist for some region of temperature. In contrast, at a second order transition the two free energy curves meet tangentially.

2.1.2.4 Phase diagrams

Phase transitions occur as one of several different thermodynamic fields is varied. Thus, the loci of all points at which phase transitions occur form phase boundaries in a multidimensional space of thermodynamic fields. The classic example of a phase diagram is that of water, shown in pressure–temperature space in Fig. 2.4, in which lines of first order transitions separate ice–water, water–steam, and ice–steam. The three first order transitions join at a 'triple point', and the water–steam phase line ends at a 'critical point' where a second

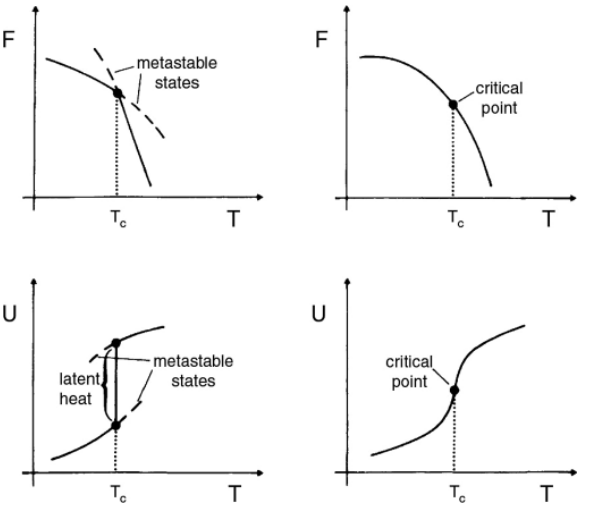


Fig. 2.3 (left) Schematic temperature dependence of the free energy and the internal energy for a system undergoing a first order transition; (right) schematic temperature dependence of the free energy and the internal energy for a system undergoing a second order transition.

order phase transition occurs. (Ice actually has multiple inequivalent phases and we have ignored this complexity in this figure.) Predicting the phase diagram of simple atomic or molecular systems, as well as of mixtures, given the knowledge of the microscopic interactions, is an important task of statistical mechanics which relies on simulation methods quite strongly, as we shall see in later chapters. A much simpler phase diagram than for water occurs for the Ising ferromagnet with Hamiltonian

$$\mathcal{H} = -\mathcal{J}_{nn} \sum_{nn} \sigma_i \sigma_j - H \sum_i \sigma_i , \tag{2.24}$$

where $\sigma_i = \pm 1$ represents a 'spin' at lattice site i which interacts with nearest neighbors on the lattice with interaction constant $\mathcal{J}_{nn} > 0$. In many respects this model has served as a 'fruit fly' system for studies in statistical mechanics. At low temperatures a first order transition occurs as H is swept through zero, and the phase boundary terminates at the critical temperature T_c as shown in Fig. 2.4. In this model it is easy to see, by invoking the symmetry involving reversal of all the spins and the sign of H, that the phase boundary must occur

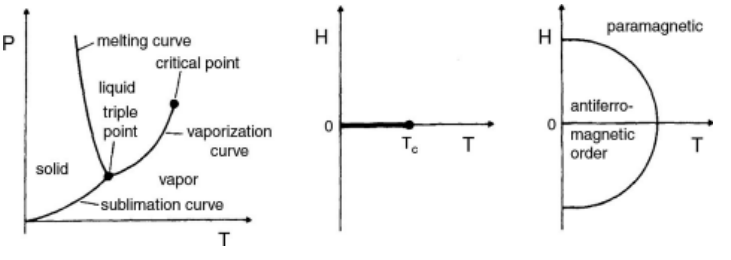


Fig. 2.4 (left) Simplified pressure–temperature phase diagram for water; (center) magnetic field–temperature phase diagram for an Ising ferromagnet; (right) magnetic field–temperature phase diagram for an Ising antiferromagnet.

at $H = 0$ so that the only remaining 'interesting' question is the location of the critical point. Of course, many physical systems do not possess this symmetry. As a third example, in Fig. 2.4 we also show the phase boundary for an Ising antiferromagnet for which $\mathcal{J} < 0$. Here the antiferromagnetic phase remains stable in non-zero field, although the critical temperature is depressed. As in the case of the ferromagnet, the phase diagram is symmetric about $H = 0$. We shall return to the question of phase diagrams for the antiferromagnet later in this section when we discuss 'multicritical points'.

2.1.2.5 Critical behavior and exponents

We give here a somewhat detailed account of critical behavior since Monte Carlo simulation is one of the best suited methods for delivering quantitative information about critical behavior. We shall attempt to explain thermodynamic singularities in terms of the reduced distance from the critical temperature. Extensive experimental research has long provided a testing ground for developing theories (Kadanoff *et al.*, 1967) and more recently, of course, computer simulations have been playing an increasingly important role. Of course, experiment is limited not only by instrumental resolution but also by unavoidable sample imperfections. Thus, the beautiful specific heat peak for $RbMnF_3$, shown in Fig. 2.5, is quite difficult to characterize for $\varepsilon \leq 10^{-4}$. Data from multiple experiments as well as results for a number of exactly soluble models show that the thermodynamic properties can be described by a set of simple power laws in the vicinity of the critical point T_c, e.g. for a magnet the order parameter m, the specific heat C, the susceptibility χ, and the correlation length ξ vary as (Stanley, 1971; Fisher, 1974)

$$m = m_o \varepsilon^{\beta}, \tag{2.25a}$$

$$\chi = \chi_o \varepsilon^{-\gamma} \tag{2.25b}$$

$$C = C_o \varepsilon^{-\alpha}, \tag{2.25c}$$

$$\xi = \xi_o \varepsilon^{-v} \tag{2.25d}$$

where $\varepsilon = |1 - T/T_c|$ and the powers (Greek characters) are termed 'critical exponents'. Note that Eqns. (2.25a–d) represent asymptotic expressions which are valid only as $\varepsilon \to 0$ and more complete forms would include additional 'corrections to scaling' terms which describe the deviations from the asymptotic behavior. Although the critical exponents for a given quantity are believed to be identical when T_c is approached from above or below, the prefactors, or 'critical amplitudes' are not usually the same. The determination of particular amplitude ratios does indeed form the basis for rather extended studies (Privman *et al.*, 1991). Along the critical isotherm, i.e. at $T = T_c$ we can define another exponent (for a ferromagnet) by

$$m = DH^{1/\delta}, \tag{2.26}$$

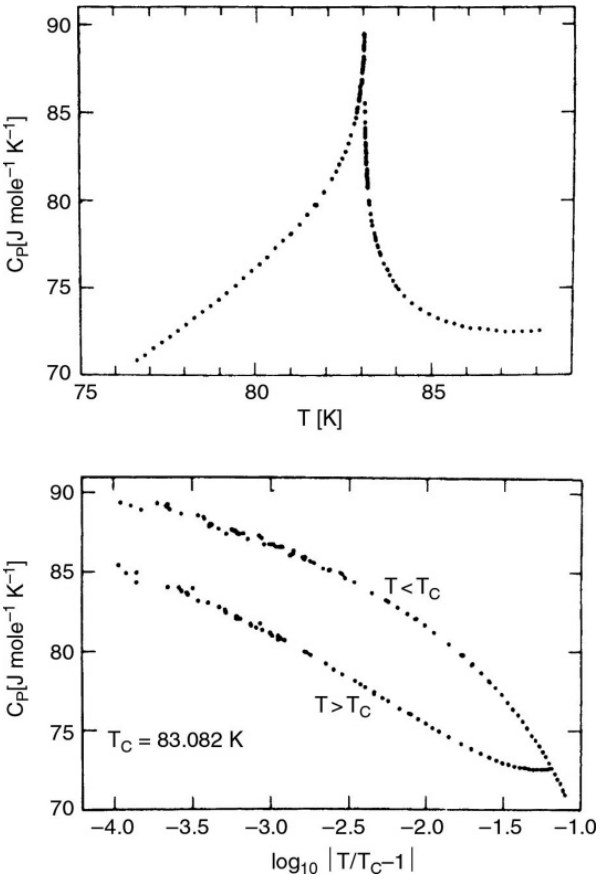


Fig. 2.5 (top) Experimental data and (bottom) analysis of the critical behavior of the specific heat of the Heisenberg-like antiferromagnet $RbMnF_3$. The critical temperature is T_c. After Kornblit and Ahlers (1973).

where H is an applied, uniform magnetic field. (Here too, an analogous expression would apply for a liquid–gas system at the critical temperature as a function of the deviation from the critical pressure.) For a system in d-dimensions the two-body correlation function $\Gamma(r)$, which well above the critical temperature has the Ornstein–Zernike form (note that for a ferromagnet in zero field $\rho(r)$ in Eqn. (2.23) corresponds to the magnetization density at r while for a fluid $\rho(r)$ means the local deviation from the average density)

$$\Gamma(r) \propto r^{-(d-1)/2} \exp(-r/\xi), \qquad r \to \infty, \tag{2.27}$$

also shows a power law decay at T_c,

$$\Gamma(r) = \Gamma_0 r^{-(d-2+\eta)}, \qquad r \to \infty, \tag{2.28}$$

where η is another critical exponent. These critical exponents are known exactly for only a small number of models, most notably the two-dimensional Ising square lattice (Onsager, 1944) (cf. Eqn. (2.24)), whose exact solution shows that $\alpha = 0$, $\beta = 1/8$, and $\gamma = 7/4$. Here, $\alpha = 0$ corresponds to a logarithmic divergence of the specific heat. We see in Fig. 2.5, however, that the experimental data for the specific heat of $RbMnF_3$ increases even more

slowly than a logarithm as $\varepsilon \to 0$, implying that $\alpha < 0$, i.e. the specific heat is non-divergent. In fact, a suitable model for $RbMnF_3$ is not the Ising model but a three-dimensional Heisenberg model with classical spins of unit length and nearest neighbor interactions

$$\mathcal{H} = -\mathcal{J} \sum_{nn} (S_{ix} S_{jx} + S_{iy} S_{jy} + S_{iz} S_{jz}), \tag{2.29}$$

which has different critical exponents than does the Ising model. (Although no exact solutions are available, quite accurate values of the exponents have been known for some time due to application of the field theoretic renormalization group (Zinn-Justin and LeGuillou, 1980), and extensive Monte Carlo simulations have yielded some rather precise results, at least for classical Heisenberg models (Chen *et al.*, 1993).)

The above picture is not complete because there are also special cases which do not fit into the above scheme. Most notable are two-dimensional XY-models with Hamiltonian

$$\mathcal{H} = -\mathcal{J} \sum_{nn} (S_{ix} S_{jx} + S_{iy} S_{jy}), \tag{2.30}$$

where S_i is a unit vector which may have either two components (plane rotator model) or three components (XY-model). These models develop no long range order at low temperature but have topological excitations, termed vortex–antivortex pairs, which unbind at the transition temperature T_{KT} (Kosterlitz and Thouless, 1973). The correlation length and susceptibility for this model diverge exponentially fast as the transition temperature is approached from above, e.g.

$$\xi \propto \exp(a \varepsilon^{-v}), \tag{2.31}$$

and every temperature below T_{KT} is a critical point. Other classical models with suitable competing interactions or lattice structures may also show 'unusual' transitions (Landau, 1994) which in some cases include different behavior of multiple order parameters at T_c and which are generally amenable to study by computer simulation.

The above discussion was confined to static aspects of phase transitions and critical phenomena. The entire question of dynamic behavior will be treated in a later section using extensions of the current formulation. A good review and comparison of renormalization group, series expansion, and Monte Carlo results for several simple 'benchmark' spin models are to be found in Pelissetto and Vicari (2002).

2.1.2.6 Universality and scaling

Homogeneity arguments also provide a way of simplifying expressions which contain thermodynamic singularities. For example, for a simple Ising ferromagnet in a small magnetic field H and at a temperature T which is near the

critical point, the singular portion of the free energy $F(T, H)$ can be written as

$$F_s = \varepsilon^{2-\alpha}\mathcal{F}^{\pm}(H/\varepsilon^{\Delta}), \tag{2.32}$$

where the 'gap exponent' Δ is equal to $\frac{1}{2}(2-\alpha+\gamma)$ and $\mathcal{F}^{\pm}$ is a function of the 'scaled' variable (H/ε^{Δ}), i.e. does not depend upon ε independently. This formula has the consequence, of course, that all other expressions for thermodynamic quantities, such as specific heat, susceptibility, etc., can be written in scaling forms as well. Similarly, the correlation function can be expressed as a scaling function of two variables

$$\Gamma(r, \xi, \varepsilon) = r^{-(d-2+\eta)}\mathcal{G}(r/\xi, H/\varepsilon^{\Delta}), \tag{2.33}$$

where $\mathcal{G}(x, y)$ is now a scaling function of two variables.

Not all of the six critical exponents defined in the previous section are independent, and using a number of thermodynamic arguments one can derive a series of exponent relations called scaling laws which show that only two exponents are generally independent. For example, taking the derivative of the free energy expressed above in a scaling form yields

$$-\partial F_s/\partial H = M = \varepsilon^{2-\alpha-\Delta}\mathcal{F}'(H/\varepsilon^{\Delta}), \tag{2.34}$$

where $\mathcal{F}'$ is the derivative of $\mathcal{F}$, but this equation can be compared directly with the expression for the decay of the order parameter to show that $\beta = 2 - \alpha - \Delta$. Furthermore, using a scaling expression for the magnetic susceptibility

$$\chi = \varepsilon^{-\gamma}\mathcal{C}(H/\varepsilon^{\Delta}) \tag{2.35}$$

one can integrate to obtain the magnetization, which for $H = 0$ becomes

$$m \propto \varepsilon^{\Delta-\gamma}. \tag{2.36}$$

Combining these simple relations one obtains the so-called Rushbrooke equality

$$\alpha + 2\beta + \gamma = 2 \tag{2.37}$$

which should be valid regardless of the individual exponent values. Another scaling law which has important consequences is the 'hyperscaling' expression which involves the lattice dimensionality d

$$d\nu = 2 - \alpha. \tag{2.38}$$

Of course, here we are neither concerned with a discussion of the physical justification of the homogeneity assumption given in Eqn. (2.32), nor with this additional scaling relation, Eqn. (2.38), see e.g. Yeomans (1992). However, these scaling relations are a prerequisite for the understanding of finite size scaling which is a basic tool in the analysis of simulational data near phase transitions, and we shall thus summarize them here. Hyperscaling may be violated in some cases, e.g. the upper critical (spatial) dimension for the Ising model is $d = 4$ beyond which mean-field (Landau theory) exponents apply and

hyperscaling is no longer obeyed. Integration of the correlation function over all spatial displacement yields the susceptibility

$$\chi \propto \varepsilon^{-\nu(2-\eta)}, \tag{2.39}$$

and by comparing this expression with the 'definition', cf. Eqn. (2.25b), of the critical behavior of the susceptibility we have

$$\gamma = \nu(2 - \eta). \tag{2.40}$$

Those systems which have the same set of critical exponents are said to belong to the same universality class (Fisher, 1974). Relevant properties which play a role in the determination of the universality class are known to include spatial dimensionality, spin dimensionality, symmetry of the ordered state, the presence of symmetry breaking fields, and the range of interaction. Thus, nearest neighbor Ising ferromagnets (see Eqn. (2.24)) on the square and triangular lattices have identical critical exponents and belong to the same universality class. Further, in those cases where lattice models and similar continuous models with the same symmetry can be compared, they generally belong to the same universality class. A simple, nearest neighbor Ising antiferromagnet in a field has the same exponents for all field values below the zero temperature critical field. This remarkable behavior will become clearer when we consider the problem in the context of renormalization group theory (Wilson, 1971) in Chapter 9. At the same time there are some simple symmetries which can be broken quite easily. For example, an isotropic ferromagnet changes from the Heisenberg universality class to the Ising class as soon as a uniaxial anisotropy is applied to the system:

$$\mathcal{H} = -\mathcal{J} \sum_{ij} [(1 - \Delta)(S_{ix} S_{jx} + S_{iy} S_{jy}) + S_{iz} S_{jz}], \tag{2.41}$$

where $\Delta > 0$. The variation of the critical temperature is then given by

$$T_c(\Delta) - T_c(\Delta = 0) \propto \Delta^{1/\phi}, \tag{2.42}$$

where ϕ is termed the 'crossover exponent' (Riedel and Wegner, 1972). There are systems for which the lattice structure and/or the presence of competing interactions give rise to behavior which is in a different universality class than one might at first believe from a cursory examination of the Hamiltonian. From an analysis of the symmetry of different possible adlayer structures for adsorbed films on crystalline substrates Domany *et al.* (1980) predict the universality classes for a number of two-dimensional Ising-lattice gas models. Among the most interesting and unusual results of this symmetry analysis is the phase diagram for the triangular lattice gas (Ising) model with nearest neighbor repulsive interaction and next-nearest neighbor attractive coupling (Landau, 1983). In the presence of non-zero chemical potential, the groundstate is a three-fold degenerate state with 1/3 or 2/3 filling (the triangular lattice splits into three sublattices and one is full and the other two are empty, or vice versa,

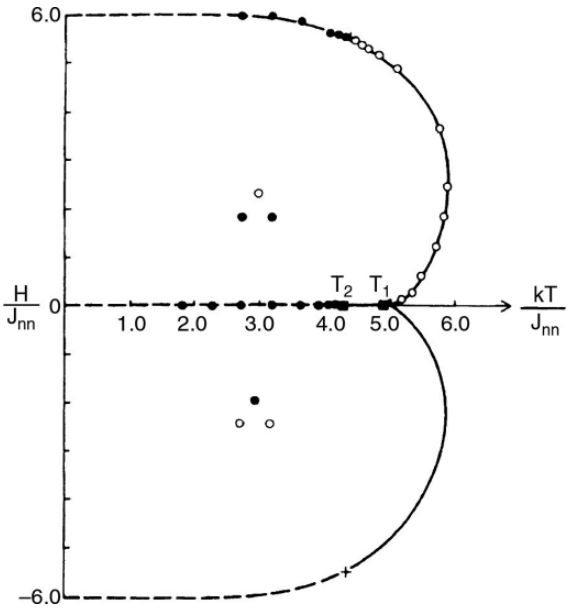


Fig. 2.6 Phase diagram for the triangular Ising (lattice gas) model with antiferromagnetic nearest neighbor and ferromagnetic next-nearest neighbor interactions. T_1 and T_2 denote Kosterlitz–Thouless phase transitions and the + sign on the non-zero field phase boundary is a tricritical point. The arrangement of open and closed circles shows examples of the two different kinds of groundstates using lattice gas language. From Landau (1983).

respectively) and is predicted to be in the universality class of the three-state Potts model (Potts, 1952; Wu, 1982)

$$\mathcal{H} = -\mathcal{J} \sum_{ij} \delta_{\sigma_i \sigma_j}, \tag{2.43}$$

where $\sigma_i = 1, 2,$ or 3. In zero chemical potential all six states become degenerate and a symmetry analysis predicts that the system is then in the universality class of the XY-model with sixth order anisotropy

$$\mathcal{H} = -\mathcal{J} \sum_{ij} (S_{ix} S_{jx} + S_{iy} S_{jy}) + \Delta \sum_{i} \cos(6\theta_i), \tag{2.44}$$

where θ_i is the angle which a spin makes with the x-axis. Monte Carlo results (Landau, 1983), shown in Fig. 2.6, confirm these expectations: in non-zero chemical potential there is a Potts-like phase boundary, complete with a three-state Potts tricritical point. (Tricritical points will be discussed in the following sub-section.) In zero field, there are two Kosterlitz–Thouless transitions with an XY-like phase separating a low temperature ordered phase from a high temperature disordered state. Between the upper and lower transitions 'vortex-like' excitations can be identified and followed. Thus, even though the Hamiltonian is that of an Ising model, there is no Ising behavior to be seen and instead a very rich scenario, complete with properties expected only for continuous spin models is found. At the same time, Fig. 2.6 is an example of a phase diagram containing both continuous and first order phase transitions which cannot yet be found with *any* other technique with an accuracy which is competitive to that obtainable by the Monte Carlo methods which will be described in this book.

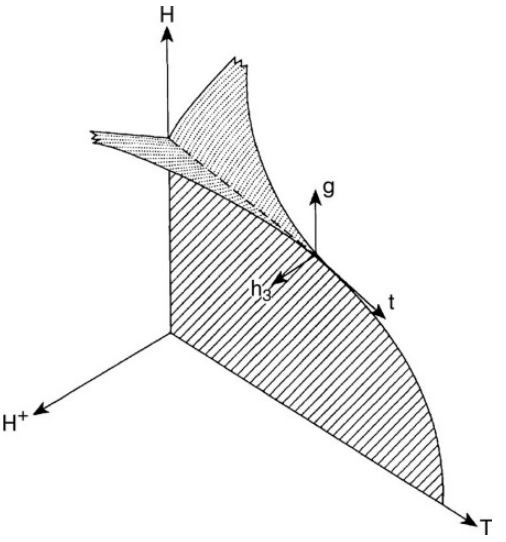


Fig. 2.7 Phase diagram for a system with a tricritical point in the three-dimensional thermodynamic field space which includes both ordering and non-ordering fields. Tricritical scaling axes are labeled t, g, and h_3.

2.1.2.7 Multicritical phenomena

Under certain circumstances the order of a phase transition changes as some thermodynamic parameter is altered. Although such behavior appears to violate the principles of universality which we have just discussed, examination of the system in a larger thermodynamic space makes such behavior easy to understand. The intersection point of multiple curves of second order phase transitions is known as a multicritical point. Examples include the tricritical point (Griffiths, 1970; Stryjewski and Giordano, 1977; Lawrie and Sarbach, 1984) which occurs in He^3-He^4 mixtures, strongly anisotropic ferromagnets, and ternary liquid mixtures, as well as the bicritical point (Nelson *et al.*, 1974) which appears on the phase boundary of a moderately anisotropic Heisenberg antiferromagnet in a uniform magnetic field. The characteristic phase diagram for a tricritical point is shown in Fig. 2.7 in which one can see that the three second order boundaries to first order surfaces of phase transitions meet at a tricritical point. One of the simplest models which exhibits such behavior is the Ising antiferromagnet with nearest and next-nearest neighbor coupling

$$\mathcal{H} = -\mathcal{J}_{nn} \sum_{nn} \sigma_i \sigma_j - \mathcal{J}_{nnn} \sum_{nnn} \sigma_i \sigma_j - H \sum_i \sigma_i - H^+ \sum_i \sigma_i, \quad (2.45)$$

where $\sigma_i = \pm 1$, H is the uniform magnetic field, and H^+ is the staggered magnetic field which couples to the order parameter. The presence of a multicritical point introduces a new 'relevant' field g, which as shown in Fig. 2.7 makes a non-zero angle with the phase boundary, and a second scaling field t, which is tangential to the phase boundary at the tricritical point. In the vicinity of a multicritical point a 'crossover' scaling law is valid

$$F(\varepsilon, H^+, g) = |g|^{2-\alpha_\varepsilon} \mathcal{F}(H^+/|g|^{\Delta_\varepsilon}, \varepsilon/|g|^{\phi_\varepsilon}), \quad (2.46)$$

where α_ε is the specific heat exponent appropriate for a tricritical point, Δ_ε the corresponding 'gap exponent', and ϕ_ε a new 'crossover' exponent. In addition, there are power law relations which describe the vanishing of discontinuities as the tricritical point is approached from below. For example, the discontinuity in the magnetization from M^+ to M^- as the first order phase boundary for $T < T_t$ is crossed decreases as

$$\Delta M = M^+ - M^- \propto |1 - T/T_t|^{\beta_u}. \tag{2.47}$$

The 'u-subscripted' exponents are related to the 'ε-subscripted' ones by a crossover exponent,

$$\beta_u = (1 - \alpha_\varepsilon)/\phi_\varepsilon. \tag{2.48}$$

As will be discussed below, the mean field values of the tricritical exponents are $\alpha_\varepsilon = 1/2$, $\Delta_\varepsilon = 5/2$, $\phi_\varepsilon = 1/2$, and hence $\beta_u = 1$. In $d = 2$ dimensions, the tricritical Ising exponents can be obtained exactly from conformal invariance methods (Nienhuis, 1982): $\alpha_\varepsilon = 8/9$, $\beta_\varepsilon = 1/24$, $\gamma_\varepsilon = 37/36$, $\nu_\varepsilon = 5/9$, $\Delta_\varepsilon = 77/72$, $\phi = 4/9$, and $\beta_u = 1/4$. We note, for comparison, that Monte Carlo renormalization group methods (see Chapter 9) had determined the following values *before* the conformal invariance results were available: $\alpha_\varepsilon = 0.89$, $\beta_\varepsilon = 0.039$, $\gamma_\varepsilon = 1.03$, $\nu_\varepsilon = 0.56$, $\Delta_\varepsilon = 1.07$, and $\phi_\varepsilon = 0.47$ (Landau and Swendsen, 1981). Tricritical points have been explored using both computer simulations of model systems as well as by experimental investigation of physical systems, and their theoretical aspects have been studied in detail (Lawrie and Sarbach, 1984).

2.1.2.8 Landau theory

One of the simplest theories with which simulations are often compared is the Landau theory, which begins with the assumption that the free energy of a system can be expanded about the phase transition in terms of the order parameter. The free energy of a d-dimensional system near a phase transition is expanded in terms of a simple one-component order parameter $m(x)$

$$F = F_o + \int d^d x \left\{ \frac{1}{2} r m^2(x) + \frac{1}{4} u m^4(x) + \frac{1}{6} v m^6(x) - \frac{H}{k_B T} m(x) + \frac{1}{2d} [R \nabla m(x)]^2 + \cdots \right\}. \tag{2.49}$$

Here a factor of $(k_B T)^{-1}$ has been absorbed into F and F_o and the coefficients r, u, and v are dimensionless. Note that the coefficient R can be interpreted as the interaction range of the model. This equation is in the form of a Taylor series in which symmetry has already been used to eliminate all odd order terms for $H = 0$. For more complex systems it is possible that additional terms, e.g. cubic products of components of a multicomponent order parameter might appear, but such situations are generally beyond the scope of our present

treatment. In the simplest possible case of a homogeneous system this equation becomes

$$F = F_o + V\left(\tfrac{1}{2}rm^2 + \tfrac{1}{4}um^4 + \tfrac{1}{6}vm^6 - mH/k_BT + \cdots\right). \quad (2.50)$$

In equilibrium the free energy must be a minimum; and if $u > 0$ we can truncate the above equation and the minimization criterion $\partial F/\partial m = 0$ yields three possible solutions:

$$m_1 = 0, \quad (2.51a)$$

$$m_{2,3} = \pm\sqrt{-r/u}. \quad (2.51b)$$

Expanding r in the vicinity of T_c so that $r = r'(T - T_c)$ we find then for $r < 0$ (i.e. $T < T_c$)

$$m_{2,3} = \pm[(r'T_c/u)(1 - T/T_c)]^{1/2}. \quad (2.52)$$

Thus, m_1 corresponds to the solution above T_c where there is no long range order, and $m_{2,3}$ correspond to solutions below T_c where the order parameter approaches zero with a characteristic power law (see Eqn. (2.25a)) with exponent $\beta = 1/2$. A similar analysis of the susceptibility produces $\gamma = 1$, $\delta = 3$. (Although straightforward to apply, Landau theory does not correctly describe the behavior of many physical systems. For liquid–gas critical points and most magnetic systems $\beta \approx 1/3$ (Kadanoff *et al.*, 1967) instead of the Landau value of $\beta = 1/2$.) The appearance of tricritical points can be easily understood from the Landau theory. If the term in m^4 is negative it becomes necessary to keep the sixth order term and the minimization process yields five solutions:

$$m_1 = 0, \quad (2.53a)$$

$$m_{2,3} = \pm\left[\frac{1}{2v}\left(-u + \sqrt{u^2 - 4rv}\right)\right]^{1/2} \quad (2.53b)$$

$$m_{4,5} = \pm\left[\frac{1}{2v}\left(-u - \sqrt{u^2 - 4rv}\right)\right]^{1/2} \quad (2.53c)$$

If v is positive, there are multiple solutions and the transition is first order. A tricritical point thus appears when $r = u = 0$, and the tricritical exponents which result from this analysis are

$$\alpha_t = \tfrac{1}{2}, \quad (2.54a)$$

$$\beta_t = \tfrac{1}{4}, \quad (2.54b)$$

$$\gamma_t = 1, \quad (2.54c)$$

$$\delta_t = 5. \quad (2.54d)$$

Note that these critical exponents are different from the values predicted for the critical point. The crossover exponent is predicted by Landau theory to be $\phi = \frac{1}{2}$.

2.1.3 Ergodicity and broken symmetry

The principle of ergodicity states that all possible configurations of the system should be attainable. As indicated in Eqn. (2.4) the different states will not all have the same probability, but it must nonetheless be possible to reach each state with non-zero probability. Below a phase transition multiple different ordered states may appear, well separated in phase space. If the phase transition from the disordered phase to the ordered phase is associated with 'symmetry breaking', the separate ordered states are related by a symmetry operation acting on the order parameter (e.g. a reversal of the sign of the order parameter for an Ising ferromagnet). In the context of a discussion of dynamical behavior of such systems, symmetry breaking usually means ergodicity breaking, i.e. the system stays in one separate region in phase space. The question of non-ergodic behavior in the context of simulations is complex. For example, in the simulation of an Ising system which may have all spins up or all spins down, we may wish to keep the system from exploring all of phase space so that only positive values of the order parameter are observed. If instead the simulation algorithm is fully ergodic, then both positive and negative values of order parameter will appear and the average will be zero. A danger for simulations is that specialized algorithms may be unintentionally non-ergodic, thus yielding incorrect results.

2.1.4 Fluctuations and the Ginzburg criterion

As mentioned earlier, the thermodynamic properties of a system are not perfectly constant but fluctuate with time as the system explores different regions of phase space. In the discussion of fluctuations in Section 2.1.1.4 we have seen that relative fluctuations of extensive thermodynamic variables scale inversely with V or N, and hence such global fluctuations vanish in the thermodynamic limit. One should not conclude, however, that fluctuations are generally unimportant; indeed local fluctuations can have dramatic consequences and require a separate discussion.

What is the importance of local fluctuations? As long as they do not play a major role, we can expect that Landau theory will yield correct predictions. Let us compare the fluctuations in $m(x)$ for a d-dimensional system over the 'correlation volume' ξ^d with its mean value m_o. If Landau theory is valid and fluctuations can be ignored, then

$$\frac{\langle [m(x) - m_o]^2 \rangle}{m_o^2} \ll 1. \tag{2.55}$$

This condition, termed the Ginzburg criterion, leads to the expression

$$\xi^d m_o^2 \chi^{-1} \gg \text{const.}, \tag{2.56}$$

and following insertion of the critical behavior power laws we obtain

$$\varepsilon^{-\nu d + 2\beta + \gamma} \gg \text{const.}, \tag{2.57}$$

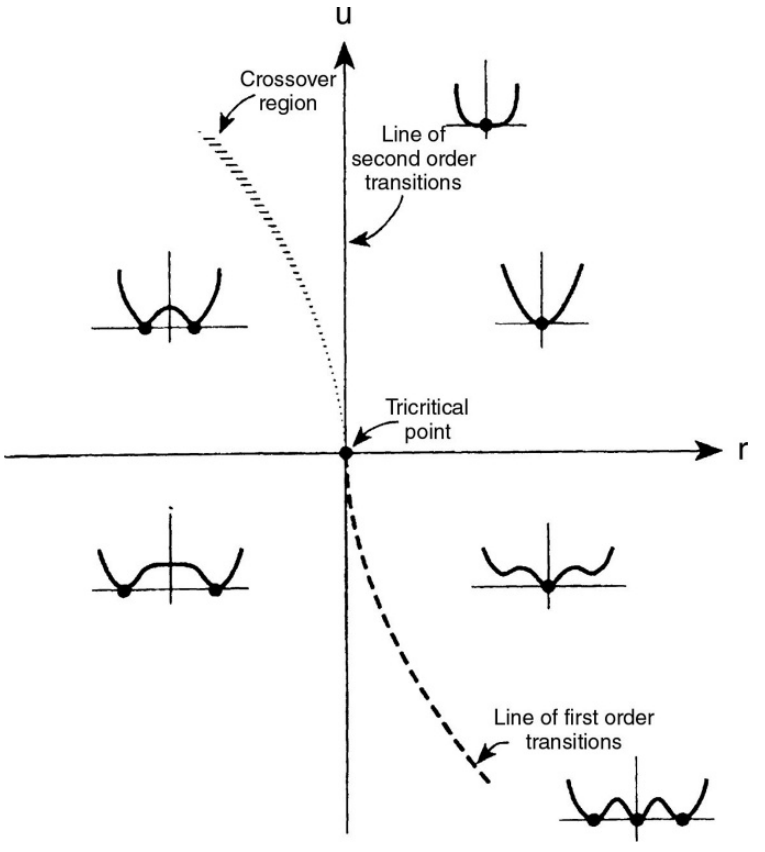


Fig. 2.8 Landau free energy and phase boundaries for the m^6 model in the $r-u$ plane. The heavy solid line shows the second order phase boundary and the dashed line represents the first order portion of the phase boundary. The heavy dots show the location(s) of the minimum free energy.

Inserting Landau exponents into this expression we find

$$\varepsilon^{(d-4)/2} \ll \text{const.}, \tag{2.58}$$

i.e. for Landau theory to be valid the lattice dimensionality must be greater than or equal to the upper critical dimension $d_u = 4$. In addition, below some lower critical dimensionality d_l fluctuations dominate completely and no transition occurs. In order to consider the tricritical point scenario depicted in Fig. 2.7, it becomes necessary to retain the next order term $\sim vm^6$ in the Landau free energy. The shape of the resultant free energy is shown in Fig. 2.8 below, at and above the tricritical point. It turns out that mean field (i.e. Landau) theory is valid for tricritical behavior above an upper critical dimension; for the Ising model with competing interactions $d_u = 3$, but for $d = 3$ there are logarithmic corrections (Wegner and Riedel, 1973).

2.1.5 A standard exercise: the ferromagnetic Ising model

The Ising model of magnetism, defined in Eqn. (2.24), is extremely well suited to Monte Carlo simulation. The same model is equivalent to simple lattice gas models for liquid–gas transitions or binary alloy models. The transformation to

a lattice gas model is straightforward. We first define site occupation variables c_i which are equal to l if the site is occupied and 0 if the site is empty. These variables are simply related to the Ising variables by

$$c_i = (1 + \sigma_i)/2. \tag{2.59}$$

If we now substitute these into the Ising Hamiltonian we find

$$\mathcal{H}_{1g} = -\phi \sum_{ij} c_i c_j - \mu \sum_i c_i + \text{const.}, \tag{2.60}$$

where $\phi = 4\mathcal{J}$ and $\mu = 2(H + 4z\mathcal{J})$ if there are z interacting neighbors. Note that if the Ising model is studied in the canonical ensemble, any spin-flips change the number of particles in the lattice gas language and the system is effectively being studied in the grand canonical ensemble. A Monte Carlo program follows a stochastic path through phase space, a procedure which will be discussed in detail in the following chapters, yielding a sequence of states from which mean values of system properties may be determined. In the following example we show what a sample output from a Monte Carlo run might look like. A complete description of the simulation algorithm, methods of analysis, and error determination will be discussed in Chapter 4.

Example

Sample output from a Monte Carlo program simulating the two-dimensional Ising model ($\mathcal{J} = 1$) at $k_B T = 1.5$ for $L = 6$, with periodic boundary conditions.

1000 MCS discarded for equilibration
5000 MCS retained for averages
1000 MCS per bin

bin	$E(t)$	$M(t)$
1	−1.9512	0.9866
2	−1.9540	0.9873
3	−1.9529	0.9867
4	−1.9557	0.9878
5	−1.9460	0.9850

Averages: $\langle E \rangle = -1.952 \pm 0.026$
$\langle M \rangle = 0.987 \pm 0.014$

specific heat = 0.202
susceptibility = 0.027

final state:

```
+ + + + + +
+ − + + + +
+ + + + + +
+ + + + − +
+ + − + + +
+ + + + + +
```

Problem 2.3 Use the fluctuation relation for the magnetization together with Eqn. (2.59) to derive a fluctuation relation for the particle number in the grand canonical ensemble of the lattice gas.

2.2 PROBABILITY THEORY

2.2.1 Basic notions

It will soon become obvious that the notions of probability and statistics are essential to statistical mechanics and, in particular, to Monte Carlo simulations in statistical physics. In this section we want to remind the reader about some fundamentals of probability theory. We shall restrict ourselves to the basics; far more detailed descriptions may be found elsewhere, for example in the books by Feller (1968) or Kalos and Whitlock (1986). We begin by considering an elementary event with a countable set of random outcomes, $A_1, A_2, \ldots, A_k$ (e.g. rolling a die). Suppose this event occurs repeatedly, say N times, with $N \gg 1$, and we count how often the outcome A_k is observed (N_k). Then it makes sense to define probabilities p_k for the outcome A_k or (we assume that all possible events have been enumerated)

$$p_k = \lim_{N\to\infty} (N_k/N), \qquad \sum_k p_k = 1. \tag{2.61}$$

Obviously we have $0 \leq p_k \leq 1$ (if A_k never occurs, $p_k = 0$; if it is certain to occur, $p_k = 1$). An equivalent notation, convenient for our purposes, is $P(A_k) \equiv p_k$. From its definition, we conclude that $P(A_i \text{ and/or } A_j) \leq [P(A_i) + P(A_j)]$. We call A_i and A_j 'mutually exclusive' events, if, and only if, the occurrence of A_i implies that A_j does not occur and vice versa. Then

$$P(A_i \text{ and } A_j) = 0, \qquad P(A_i \text{ or } A_j) = P(A_i) + P(A_j). \tag{2.62}$$

Let us now consider two events, one with outcomes $\{A_i\}$ and probabilities p_{1i}; the second with outcomes $\{B_j\}$ and probabilities p_{2j}, respectively. We consider now the outcome $(A_i; B_j)$ and define p_{ij} as the joint probability that both A_i and B_j occur. If the events are independent, we have

$$p_{ij} = p_{1i} \times p_{2j}. \tag{2.63}$$

If they are not independent, it makes sense to define the conditional probability $p(j|i)$ that B_j occurs, given that A_i occurs

$$p(j|i) = \frac{p_{ij}}{\sum_k p_{ik}} = \frac{p_{ij}}{p_{1i}}. \tag{2.64}$$

Of course we have $\sum_j p(j|i) = 1$ since some B_j must occur.

The outcome of such random events may be logical variables (True or False) or real numbers x_i. We call these numbers random variables. We now define

the expectation value of this random variable as follows:

$$\langle x \rangle \equiv E(x) \equiv \sum_i p_i x_i. \tag{2.65}$$

Similarly, any (real) function $g(x_i)$ then has the expectation value

$$\langle g(x) \rangle \equiv E(g(x)) = \sum_i p_i g(x_i). \tag{2.66}$$

In particular, if we begin with two functions $g_1(x)$, $g_2(x)$ and consider the linear combination (λ_1, λ_2 being constants), we have $\langle \lambda_1 g_1(x) + \lambda_2 g_2(x) \rangle = \lambda_1 \langle g_1 \rangle + \lambda_2 \langle g_2 \rangle$. Of particular interest are the powers of x. Defining the nth moment as

$$\langle x^n \rangle = \sum_i p_i x_i^n \tag{2.67}$$

we then consider the so-called cumulants

$$\langle (x - \langle x \rangle^n \rangle = \sum_i p_i (x_i - \langle x \rangle)^n. \tag{2.68}$$

Of greatest importance is the case $n = 2$, which is called the 'variance',

$$\mathrm{var}(x) = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2. \tag{2.69}$$

If we generalize these definitions to two random variables (x_i and y_j), the analog of Eqn. (2.65) is

$$\langle xy \rangle = \sum_{i,j} p_{ij} x_i y_j. \tag{2.70}$$

If x and y are independent, then $p_{ij} = p_{1i} p_{2j}$ and hence

$$\langle xy \rangle = \sum_i p_{1i} x_i \sum_j p_{2j} y_j = \langle x \rangle \langle y \rangle. \tag{2.71}$$

As a measure of the degree of independence of the two random variables it is hence natural to take their covariance

$$\mathrm{cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle. \tag{2.72}$$

2.2.2 Special probability distributions and the central limit theorem

Do we find any special behavior which arises when we consider a very large number of events? Consider two events A_0 and A_1 that are mutually exclusive and exhaustive:

$$p(A_1) = p, \qquad x = 1; \qquad P(A_0) = 1 - p, \qquad x = 0. \tag{2.73}$$

Suppose now that N independent samples of these events occur. Each outcome is either 0 or 1, and we denote the sum X of these outcomes, $X = \sum_{r=1} x_r$.

Now the probability that $X = n$ is the probability that n of the X_r were 1 and $(N - n)$ were 0. This is called the binomial distribution,

$$P(X = n) = \binom{N}{n} p^n (1 - p)^{N-n}, \tag{2.74}$$

$\binom{N}{n}$ being the binomial coefficients. It is easy to show from Eqn. (2.74) that

$$\langle X \rangle = Np, \qquad \langle (\mathrm{X} - \langle X \rangle)^2 \rangle = Np(1 - p). \tag{2.75}$$

Suppose now we still have two outcomes (1, 0) of an experiment: if the outcome is 0, the experiment is repeated, otherwise we stop. Now the random variable of interest is the number n of experiments until we get the outcome 1:

$$P(x = n) = (1 - p)^{n-1} p, \qquad n = 1, 2, 3, \ldots \tag{2.76}$$

This is called the geometrical distribution. In the case that the probability of 'success' is very small, the Poisson distribution

$$P(x = n) = \frac{\lambda^n}{n!} \exp(-\lambda), \qquad n = 0, 1, \ldots \tag{2.77}$$

represents an approximation to the binomial distribution. The most important distribution that we will encounter in statistical analysis of data is the Gaussian distribution

$$p_{\mathrm{G}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right] \tag{2.78}$$

which is an approximation to the binomial distribution in the case of a very large number of possible outcomes and a very large number of samples. If random variables $x_1, x_2, \ldots, x_n$ are all independent of each other and drawn from the same distribution, the average value $\overline{X}_N = \sum_{i=1}^{N} x_i / N$ in the limit $N \to \infty$ will always be distributed according to Eqn. (2.78), irrespective of the distribution from which the x_i were drawn. This behavior is known as the 'central limit theorem' and plays a *very* important role in the sampling of states of a system. One also can show that the variance of $\overline{X}_N$ is the quantity σ^2 that appears in Eqn. (2.78), and that $\sigma^2 \propto 1/N$.

Of course, at this point it should be clear to those unfamiliar with probability theory that there is no way to fully understand this subject from this 'crash course' of only a few pages which we are presenting here. For the uninitiated, our goal is only to 'whet the appetite' about this subject since it is central to the estimation of errors in the simulation results. (This discussion may then also serve to present a guide to the most pertinent literature.)

Problem 2.4 Compute the average value and the variance for the exponential distribution and for the Poisson distribution.

2.2.3 Statistical errors

Suppose the quantity A is distributed according to a Gaussian with mean value $\langle A \rangle$ and width σ. We consider n statistically independent observations $\{A_i\}$ of this quantity A. An unbiased estimator of the mean $\langle A \rangle$ of this distribution is

$$\overline{A} = \frac{1}{n} \sum_{i=1}^{n} A_i \tag{2.79}$$

and the standard error of this estimate is

$$\text{error} = \sigma / \sqrt{n}. \tag{2.80}$$

In order to estimate the variance σ itself from the observations, consider deviations $\delta A_i = A_i - \overline{A}$. Trivially we have $\overline{\delta A_i} = 0$ and $\langle \delta A \rangle = 0$. Thus we are interested in the mean square deviation

$$\overline{\delta A^2} = \frac{1}{n} \sum_{i=1}^{n} (\delta A_i)^2 = \overline{A^2} - (\overline{A})^2. \tag{2.81}$$

The expectation value of this quantity is easily related to $\sigma^2 = \langle A^2 \rangle - \langle A \rangle^2$ as

$$\left\langle \overline{\delta A^2} \right\rangle = \sigma^2 (1 - 1/n). \tag{2.82}$$

Combining Eqns. (2.80) and (2.81) we recognize the usual formula for the computation of errors of averages from uncorrelated estimates,

$$\text{error} = \sqrt{\left\langle \overline{\delta A^2} \right\rangle \Big/ (n-1)} = \sqrt{\sum_{i=1}^{n} (\delta A_i)^2 \Big/ [n(n-1)]}. \tag{2.83}$$

Equation (2.83) is immediately applicable to simple sampling Monte Carlo methods. However, as we shall see later, the usual form of Monte Carlo sampling, namely importance sampling Monte Carlo, leads to 'dynamic' correlations between subsequently generated observations $\{A_i\}$. Then Eqn. (2.83) is replaced by

$$(\text{error})^2 = \frac{\sigma^2}{n} (1 + 2\tau_A / \delta t), \tag{2.84}$$

where δt is the 'time interval' between subsequently generated states A_i, A_{i+1} and τ_A is the 'correlation time' (measured in the same units as δt).

2.2.4 Markov chains and master equations

The concept of Markov chains is so central to Monte Carlo simulations that we wish to present at least a brief discussion of the basic ideas. We define a stochastic process at discrete times labeled consecutively t_1, t_2, t_3, ... for a

system with a finite set of possible states $S_1, S_2, S_3, \ldots$ and we denote by X_t the state the system is in at time t. We consider the conditional probability that $X_{t_n} = S_{i_n}$,

$$P(X_{t_n} = S_{i_n} | X_{t_{n-1}} = S_{i_{n-1}}, X_{t_{n-2}} = S_{i_{n-2}}, \ldots, X_{t_1} = S_{i_1}), \tag{2.85}$$

given that at the preceding time the system state $X_{t_{n-1}}$ was in state $S_{i_{n-1}}$, etc. Such a process is called a Markov process, if this conditional probability is in fact independent of all states but the immediate predecessor, i.e. $P = P(X_{t_n} = S_{i_n} | X_{t_{n-1}} = S_{i_{n-1}})$. The corresponding sequence of states $\{X_t\}$ is called a Markov chain, and the above conditional probability can be interpreted as the transition probability to move from state i to state j,

$$W_{ij} = W(S_i \to S_j) = P(X_{t_n} = S_j | X_{t_{n-1}} = S_i). \tag{2.86}$$

We further require that

$$W_{ij} \geq 0, \qquad \sum_j W_{ij} = 1, \tag{2.87}$$

as usual for transition probabilities. We may then construct the total probability $P(X_{t_n} = S_j)$ that at time t_n the system is in state S_j as $P(X_{t_n} = S_j) = P(X_{t_n} = S_j | X_{t_{n-1}} = S_i) P(X_{t_{n-1}} = S_i) = W_{ij} P(X_{t_{n-1}} = S_i)$.

The master equation considers the change of this probability with time t (treating time as a continuous rather than discrete variable and writing then $P(X_{t_n} = S_j) = P(S_j, t)$)

$$\frac{dP(S_j, t)}{dt} = -\sum_i W_{ji} P(S_j, t) + \sum_i W_{ij} P(S_i, t). \tag{2.88}$$

Equation (2.88) can be considered as a 'continuity equation', expressing the fact that the total probability is conserved ($\sum_j P(S_j, t) \equiv 1$ at all times) and all probability of a state i that is 'lost' by transitions to state j is gained in the probability of that state, and vice versa. Equation (2.88) just describes the balance of gain and loss processes: since the probabilities of the events $S_j \to S_{i_1}, S_j \to S_{i_2}, S_j \to S_{i_3}$ are mutually exclusive, the total probability for a move away from the state j simply is the sum $\sum_i W_{ij} P(S_j, t)$.

Of course, by these remarks we only wish to make the master equation plausible to the reader, rather than dwelling on more formal derivations. Clearly, Eqn. (2.88) brings out the basic property of Markov processes: i.e. knowledge of the state at time t completely determines the future time evolution, there is no memory of the past (knowledge of behavior of the systems at times earlier than t is not needed). This property is obviously rather special, and only some real systems actually do have a physical dynamics compatible with Eqn. (2.88), see Section 2.3.1. But the main significance of Eqn. (2.88) is that the importance sampling Monte Carlo process (like the Metropolis algorithm which will

be described in Chapter 4) can be interpreted as a Markov process, with a particular choice of transition probabilities: one must satisfy the principle of detailed balance with the equilibrium probability $P_{\rm eq}(S_j)$,

$$W_{ji} P_{\rm eq}(S_j) = W_{ij} P_{\rm eq}(S_i), \tag{2.89}$$

as will be discussed later. At this point, we already note that the master equation yields

$$dP_{\rm eq}(S_j, t)/dt \equiv 0, \tag{2.90}$$

since Eqn. (2.89) ensures that gain and loss terms in Eqn. (2.88) cancel exactly.

Finally we mention that the restriction to a discrete set of states $\{S_i\}$ is not at all important – one can generalize the discussion to a continuum of states, working with suitable probability densities in the appropriate space.

2.2.5 The 'art' of random number generation

2.2.5.1 Background

Monte Carlo methods are heavily dependent on the fast, efficient production of streams of random numbers. Since physical processes, such as white noise generation from electrical circuits, generally introduce new numbers much too slowly to be effective with today's digital computers, random number sequences are produced directly on the computer using software (Knuth, 1969). (The use of tables of random numbers is also impractical because of the huge number of random numbers now needed for most simulations and the slow access time to secondary storage media.) Since such algorithms are actually deterministic, the random number sequences which are thus produced are only 'pseudo-random' and do indeed have limitations which need to be understood. In fact, from the point of view of rigorous mathematics the use of 'pseudo-random' numbers may seem undesirable (referring to a quotation of one of the 'fathers' of computational science, John von Neumann (1951): 'Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin'), but it is inevitable. Thus, in the remainder of this book, when we refer to 'random numbers' it must be understood that we are really speaking of 'pseudo-random' numbers. These deterministic features are not always negative. For example, for testing a program it is often useful to compare the results with a previous run made using exactly the same random numbers. The explosive growth in the use of Monte Carlo simulations in diverse areas of physics has prompted extensive investigation of new methods and of the reliability of both old and new techniques. Monte Carlo simulations are subject to both statistical and systematic errors from multiple sources, some of which are well understood (Ferrenberg *et al.*, 1991). It has long been known that poor quality random number generation can lead to systematic

errors in Monte Carlo simulation (Marsaglia, 1968; Barber *et al.*, 1985); in fact, early problems with popular generators led to the development of improved methods for producing pseudo-random numbers. For an instructive analysis of the suitability of different random number generators see Coddington (1994). Useful test suites have been developed with the sole purpose of testing random number generators. Some of these can be freely downloaded from the internet, such as the NIST test sites (http://csrc.nist.gov/groups/ST/toolkit/rng) which includes a detailed documentation of these tests. We also draw attention to the Test U01 suite provided by L'Ecuyer and Simard (2007). As we shall show in the following discussion both the testing as well as the generation of random numbers remain important problems that have not been fully solved. In general, the random number sequences which are needed should be uniform, uncorrelated, and of extremely long period, i.e. do not repeat over quite long intervals. Later in this chapter we shall give some guidance on the testing for these 'desirable' properties. For a much more detailed account of random number generators, see Gentle (2003).

In the following sub-sections we shall discuss several different kinds of generators. The reason for this is that it is now clear that for optimum performance and accuracy, the random number generator needs to be matched to the algorithm and computer. Indeed, the resolution of Monte Carlo studies has now advanced to the point where *no* generator can be considered to be completely 'safe' for use with a new simulation algorithm on a new problem. The practitioner is now faced anew with the challenge of testing the random number generator for each high resolution application, and we shall review some of the 'tests' later in this section. The generators which are discussed in the next sub-sections produce a sequence of random integers. Usually floating point numbers between 0 and 1 are needed; these are obtained by carrying out a floating point divide by the largest integer $N_{\max}$ which can fit into a word.

One important topic which we shall not consider here is the question of the implementation of random number generators on massively parallel computers. In such cases one must be certain that the random number sequences on all processors are distinct and uncorrelated. As the number of processors available to single users increases, this question must surely be addressed, but we feel that at the present time this is a rather specialized topic and we shall not consider it further. This problem is particularly acute in simulations on graphics processing units (GPUs), where the number of parallel threads for Monte Carlo in a multiple-GPU simulation may be in the millions. We refer to Manssen *et al.* (2012) for a recent discussion of this problem.

2.2.5.2 Congruential method

A simple and very popular method for generating random number sequences is the multiplicative or congruential method. Here, a fixed multiplier c is chosen

along with a given seed and subsequent numbers are generated by simple multiplication:

$$X_n = (c \times X_{n-1} + a_o)\text{MOD}\, N_{\max}, \tag{2.91}$$

where X_n is an integer between 1 and $N_{\max}$. It is important that the value of the multiplier be chosen to have 'good' properties and various choices have been used in the past. In addition, the best performance is obtained when the initial random number X_0 is odd. Experience has shown that a 'good' congruential generator is the 32-bit linear congruential algorithm (CONG)

$$X_n = (16807 \times X_{n-1})\text{MOD}(2^{31} - 1). \tag{2.92}$$

A congruential generator which was quite popular earlier turned out to have quite noticeable correlation between consecutive triplets of random numbers. Nonetheless for many uses congruential generators are acceptable and are certainly easy to implement. (Congruential generators which use a longer word length also have improved properties.)

2.2.5.3 Mixed congruential methods

Congruential generators can be mixed in several ways to attempt to improve the quality of the random numbers which are produced. One simple and relatively effective method is to use two distinct generators simultaneously: the first one generates a table of random numbers and the second generator draws randomly from this table. For best results the two generators should have different seeds and different multipliers. A variation of this approach for algorithms which need multiple random numbers for different portions of the calculations is to use independent generators for different portions of the problem.

2.2.5.4 Shift register algorithms

A fast method which was introduced to eliminate some of the problems with correlations which had been discovered with a congruential method is the shift register or Tausworthe algorithm (Kirkpatrick and Stoll, 1981). A table of random numbers is first produced and a new random number is produced by combining two different existing numbers from the table:

$$X_n = X_{n-p} \cdot \text{XOR} \cdot X_{n-q}, \tag{2.93}$$

where p and q must be properly chosen if the sequence is to have good properties. The ·XOR· operator is the bitwise *exclusive-OR* operator. The best choices of the pairs (p, q) are determined by the primitive trinomials given by

$$X^p + X^q + 1 = \text{primitive}. \tag{2.94}$$

Examples of pairs which satisfy this condition are:

$$\begin{array}{ll} p = 98 & q = 27, \\ p = 250 & q = 103, \\ p = 1279 & q = 216, 418, \\ p = 9689 & q = 84, 471, 1836, 2444, 4187. \end{array}$$

R250, for which $p = 250$, $q = 103$, has been the most commonly used generator in this class. In the literature one will find cases where X_{n-q} is used and others where X_{n-p-q} is used instead. In fact, these two choices will give the same stream of numbers but in reverse order; the quality of each sequence is thus the same. In general, higher quality of random number sequences results when large values of p and q are used, although for many purposes R250 works quite well. In order for the quality of the random number sequence to be of the highest possible quality, it is important for the 'table' to be properly initialized. One simple method is to use a good congruential generator to generate the initial values; the best procedure is to use a different random number to determine each bit in succession for each entry in the initial table.

While R250 had passed all statistical tests and for a while became 'the gold standard' random number generator for applications in statistical physics, it actually works badly when it is used for simulating the two-dimensional Ising model at criticality with cluster flipping. (This will be discussed in a later chapter, see Section 5.9.7). Now, R250 has actually been referred to (Katzgraber, 2011) as a standard example for a bad random number generator.

2.2.5.5 Lagged Fibonacci generators

The shift-register algorithm is a special case of a more general class of generators known as lagged Fibonacci generators. Additional generators may be produced by replacing the exclusive-or (·XOR·) in Eqn. (2.93) by some other operator. One generator which has been found to have good properties uses the multiplication operator:

$$X_n = X_{n-p} * X_{n-q} \tag{2.95}$$

with rather small values of the 'off-set', e.g. $p = 17$, $q = 5$. More complex generators have also been used, e.g. a 'subtract with carry generator' (SWC) (Marsaglia *et al.*, 1990), which for 32-bit arithmetic is

$$\begin{aligned} X_n &= X_{n-22} - X_{n-43} - C \\ &\quad \text{if } X_n \geq 0, \qquad C = 0 \\ &\quad \text{if } X_n < 0, \qquad X_n = X_n + (2^{32} - 5), \qquad C = 1, \end{aligned} \tag{2.96}$$

and the compound generator, a combined subtract with carry-Weyl generator (SWCW) (Marsaglia *et al.*, 1990)

$$
\begin{aligned}
Z_n &= Z_{n-22} - Z_{n-43} - C \\
&\quad \text{if } Z_n \geq 0, \quad C = 0 \\
&\quad \text{if } Z_n < 0, \quad Z_n = Z_n + (2^{32} - 5), \qquad C = 1 \qquad (2.97) \\
Y_n &= (Y_{n-1} - 362436069) \text{ MOD } 2^{32} \\
X_n &= (Z_n - Y_n) \text{ MOD } 2^{32}.
\end{aligned}
$$

As mentioned earlier, it is known that the performance of a random number generator can be adversely affected by improper initialization of its lookup table (Kirkpatrick and Stoll, 1981) and we recommend the same initialization procedure for all generators as that described for R250. The above are only examples of a few different random number generators.

2.2.5.6 Tests for quality

Properties of random number generators have been carefully examined using a battery of mathematical tests (Marsaglia, 1968, 1985, unpublished); a few simple examples of such tests are:

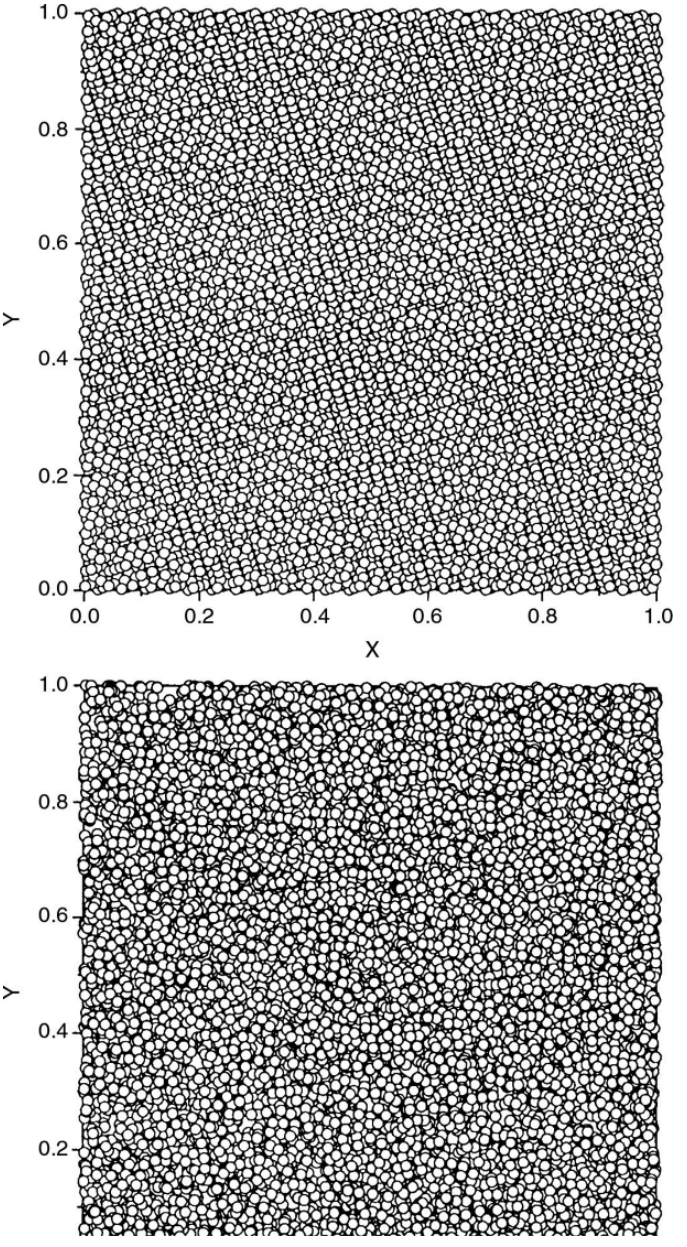
Uniformity test: Break up the interval between zero and one into a large number of small bins and after generating a large number of random numbers check for uniformity in the number of entries in each bin.
Overlapping M-tuple test: Check the statistical properties of the number of times M-tuples of digits appear in the sequence of random numbers.
Parking lot test: Plot points in an m-dimensional space where the m-coordinates of each point are determined by m-successive calls to the random number generator. Then look for regular structures.

Although the 'quality' of a sequence of random numbers is notoriously difficult to assess, often all indications from standard tests are that any residual errors from random number generation should now be smaller than statistical errors in Monte Carlo studies. However, these mathematical tests are not necessarily sufficient, and an example of a 'practical' test in a Monte Carlo study of a small lattice Ising model (which can be solved exactly) will be presented later; here both 'local' and 'non-local' sampling methods were shown to yield different levels of systematic error with different 'good' generators. (The exact nature of these algorithms is not really important at this stage and will be discussed in detail in later sections.) More sophisticated, high quality generators, such as RANLUX (James, 1994; Lüscher, 1994) which is based upon an algorithm by Marsaglia and Zaman (1991), are finding their way into use, but they are slow and must still be carefully tested with new algorithms as they are devised. (RANLUX includes two lags, plus a carry, plus it discards portions of the sequence of generated numbers. The complications tend to destroy short time correlations but have the negative effect of slowing down the generator.)

Problem 2.5 Suppose we have a computer with 4 bit words. Produce a sequence of random numbers using a congruential generator. What is the cycle length for this generator?

Example

Carry out a 'parking lot' test on two different random number generators. 10 000 points are plotted using consecutive pairs of random numbers as x- and y-coordinates. At the top is a picture of a 'bad' generator (exhibiting a striped pattern) and at the bottom are the results of a 'good' generator.



2.2.5.7 Non-uniform distributions

There are some situations in which random numbers x_i which have different distributions, e.g. Gaussian, are required. The most general way to perform this is to look at the integrated distribution function $F(x)$ of the desired distribution $f(x)$, generate a uniform distribution of random numbers y_i and then take the inverse function with the uniformly chosen random number as the variable, i.e.

$$y = F(y) = \int_0^y f(x)dx \tag{2.98}$$

So that

$$x = F^{-1}(y). \tag{2.99}$$

Example

Suppose we wish to generate a set of random numbers distributed according to $f(x) = x$. The cumulative distribution function is $y = F(x) = \int_0^x x'dx' = 0.5x^2$. If a random number y is chosen from a uniform distribution, then the desired random number is $x = 2.0y^{1/2}$.

An effective way to generate numbers according to a Gaussian distribution is the Box–Muller method. Here two different numbers x_1 and x_2 are drawn from a uniform distribution and then the desired random numbers are computed from

$$y_1 = (-2\ln x_1)^{1/2}\cos(2\pi x_2), \tag{2.100a}$$
$$y_2 = (-2\ln x_1)^{1/2}\sin(2\pi x_2). \tag{2.100b}$$

Obviously the quality of the random numbers produced depends on the quality of the uniform sequence which is generated first. Because of the extra CPU time needed for the computation of the trigonometric functions, the speed with which x_1 and x_2 are generated is not particularly important.

Problem 2.6 Given a sequence of uniformly distributed random numbers y_i, show how a sequence x_i distributed according to x^2 would be produced.

2.3 NON-EQUILIBRIUM AND DYNAMICS: SOME INTRODUCTORY COMMENTS

2.3.1 Physical applications of master equations

In classical statistical mechanics of many-body systems, dynamical properties are controlled by Newton's equations of motion for the coordinates $\boldsymbol{r}_i$ of the atoms labeled by index i, $m_i\ddot{\boldsymbol{r}}_i = -\nabla_i U$, m_i being the mass of the ith particle, and U being the total potential energy (which may contain both an external potential and interatomic contributions). The probability of a point in phase

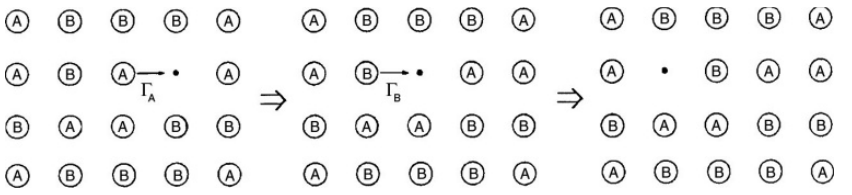


Fig. 2.9 Schematic description of interdiffusion in a model of random binary alloy (AB) with a small volume fraction of vacancies. Interdiffusion proceeds via the vacancy mechanism: A-atoms jump with rate Γ_A and B-atoms with rate Γ_B.

space then develops according to Liouville's equation, and obviously the deterministic trajectory through phase space generated in this way has nothing to do, in general, with the probabilistic trajectories generated in stochastic processes, such as Markov processes (Section 2.2.4).

However, often one is not aiming at a fully atomistic description of a physical problem, dealing with all coordinates and momenta of the atoms. Instead one is satisfied with a coarse-grained picture for which only a subset of the degrees of freedom matters. It then is rather common that the degrees of freedom that are left out (i.e. those which typically occur on a much smaller length scale and much faster time scale) act as a heat bath, inducing stochastic transitions among the relevant (and slower) degrees of freedom. In the case of a very good separation of time scales, it is in fact possible to reduce the Liouville equation to a Markovian master equation, of the type written in Eqn. (2.88).

Rather than repeating any of the formal derivations of this result from the literature, we rather motivate this description by a typical example, namely the description of interdiffusion in solid binary alloys (AB) at low temperatures (Fig. 2.9). The solid forms a crystal lattice, and each lattice site i may be occupied by an A-atom (then the concentration variable $c_i^A = 1$, otherwise $c_i^A = 0$), by a B-atom (then $c_i^B = 1$, otherwise $c_i^B = 0$), or stay vacant. Interdiffusion then happens because A-atoms jump to a (typically nearest neighbor) vacant site, with a jump rate Γ_A, and B-atoms jump to a vacant site at jump rate Γ_B, and many such random hopping events relax any concentration gradients. The distribution of the atoms over the available sites may be completely random or correlated, and the jump rates may depend on the local neighborhood or may simply be constants, etc. Now a consideration of the potential energy in solids shows that such jump events are normally thermally activated processes, $\Gamma_{A,B} \propto \exp(-\Delta E / k_B T)$, where the energy barrier to be overcome is much higher than the thermal energy (e.g. $\Delta E \approx 1\,\text{eV}$). As a result, the time a vacancy needs in order to move from one lattice site to the next one is orders of magnitude larger than the time constant of the lattice vibrations. This separation of time scales (a phonon vibration time may be of the order of 10^{-13} seconds, the time between the moves of a vacancy can be 10 orders of magnitude larger) is due to the different length scales of these motions (vibrations take only one percent of a lattice spacing at low temperatures). Thus a simulation of the dynamics of these hopping processes using the molecular dynamics method which numerically integrates Newton's equations of motion, would

suffer from a sampling of extremely rare events. The master equation, Eqn. (2.88), which can be straightforwardly simulated by Monte Carlo methods, allows the direct simulation of the important hopping events, completely disregarding the phonons. But it is also clear, of course, that knowledge of the basic rate constants for the slow degrees of freedom (the jump rates Γ_A, Γ_B in the case of our example) are an 'input' to the Monte Carlo simulation, rather than an 'output': the notion of 'time' for a Markov process (Section 2.2.4) does not specify anything about the *units* of this time. These units are only fixed if the connection between the slow degrees of freedom and the fast ones is explicitly considered, which usually is a separate problem and out of consideration here.

Although the conditions under which a master equation description of a physical system is appropriate may seem rather restrictive, it will become apparent later in this book that there is nevertheless a rich variety of physical systems and/or processes that can be faithfully modeled by this stochastic dynamics. (Examples include relaxation of the magnetization in spin glasses; Brownian motion of macromolecules in melts; spinodal decomposition in mixtures; growth of ordered monolayer domains at surfaces; epitaxial growth of multilayers; etc.)

2.3.2 Conservation laws and their consequences

Different situations may be examined in which different properties of the system are held constant. One interesting case is one in which the total magnetization of a system is conserved (held constant); when a system undergoes a first order transition it will divide into different regions in which one phase or the other dominates. The dynamics of first order transitions is a fascinating topic with many facets (Gunton *et al.*, 1983; Binder, 1987). It is perhaps instructive to first briefly review some of the static properties of a system below the critical point; for a simple ferromagnet a first order transition is encountered when the field is swept from positive to negative. Within the context of Landau theory the behavior can be understood by looking at the magnetization isotherm shown in Fig. 2.10. The solid portions of the curve are thermodynamically stable, while the dashed portions are metastable, and the dotted portion is unstable. The endpoints of the unstable region are termed 'spinodal points' and occur at magnetizations $\pm M_{sp}$. The spinodal points occur at magnetic fields $\pm H_c$. As the magnetic field is swept, the transition occurs at $H = 0$ and the limits of the corresponding coexistence region are at $\pm M_s$. If f_{cg} is a coarse-grained free energy density, then

$$\partial^2 f_{cg}/\partial M^2 = \chi_T^{-1} \to 0 \tag{2.101}$$

at the spinodal. However, this singular behavior at the spinodal is a mean-field concept, and one must ask how this behavior is modified when statistical fluctuations are considered. A Ginzburg criterion can be developed in terms of a coarse-grained length scale L and coarse-grained volume L^d. The fluctuations

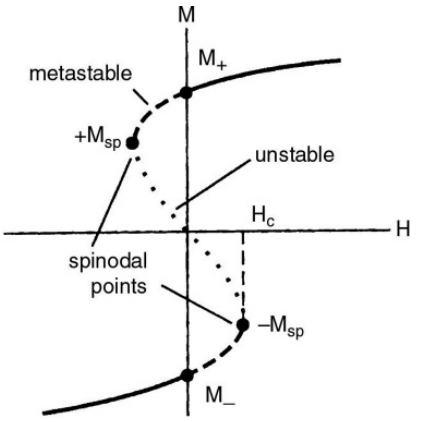


Fig. 2.10 Magnetization as a function of magnetic field for $T < T_c$. The solid curves represent stable, equilibrium regions, the dashed lines represent 'metastable', and the dotted line 'unstable' states. The values of the magnetization at the 'spinodal' are $\pm M_{sp}$ and the spinodal fields are $\pm H_c$. M_+ and M_- are the magnetizations at the opposite sides of the coexistence curve.

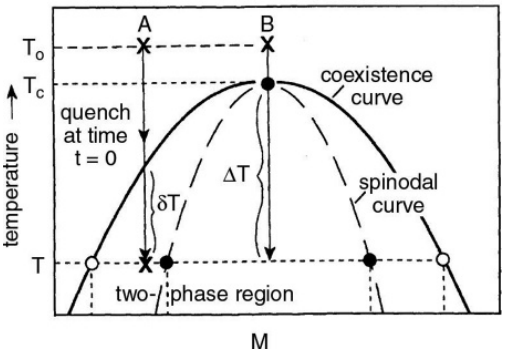


Fig. 2.11 Schematic phase coexistence diagram showing the 'spinodal' line. Paths (A) and (B) represent quenches into the nucleation regime and the spinodal decomposition regime, respectively.

in the magnetization as a function of position $M(x)$ from the mean value M must satisfy the condition

$$\langle [M(x) - M]^2 \rangle L^d / [M - M_{sp}]^2 \ll 1. \tag{2.102}$$

This leads to the condition that

$$1 \ll R^d (H_c - H)^{(6-d)/4}. \tag{2.103}$$

Thus the behavior should be mean-field-like for large interaction range R and far from the spinodal.

If a system is quenched from a disordered, high temperature state to a metastable state below the critical temperature, the system may respond in two different ways depending on where the system is immediately after the quench (see Fig. 2.11). If the quench is to a point which is close to one of the equilibrium values characteristic of the two-phase coexistence then the state evolves towards equilibrium by the nucleation and subsequent growth of 'droplets', see Fig. 2.12. (This figure is shown for pedagogical reasons and is

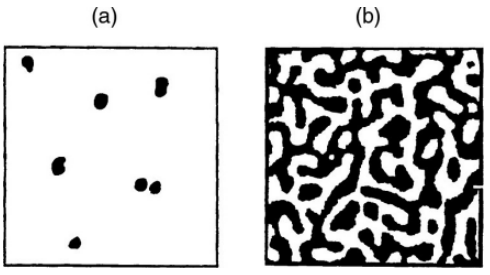


Fig. 2.12 Pictorial view of different possible modes for phase separation: (a) nucleation; (b) spinodal decomposition. The dark regions represent the phase with M_-.

not intended to provide an accurate view of the droplet formation in a particular physical system.) There will be a free energy barrier ΔF_l^* to the growth of clusters where l^* is the 'critical cluster size' and the nucleation rate $\mathcal{J}$ will be given by

$$\mathcal{J} \propto \exp(-\Delta F_l^* / k_B T). \tag{2.104}$$

Near the spinodal the argument of the exponential will be

$$\Delta F_l^* / k_B T \propto R^d (1 - T/T_c)^{4-d/2} [(M_{ms} - M_{sp})/(M_+ - M_-)]^{(6-d)/2}, \tag{2.105}$$

whereas near the coexistence curve

$$\Delta F_l^* / k_B T \propto R^d (1 - T/T_c)^{(4-d)/2} [(M_+ - M_{ms})/(M_+ - M_-)^{-(d-1)}. \tag{2.106}$$

In solid mixtures the latter stages of this growth are thought to be described by the Lifshitz–Slyozov theory (Lifshitz and Slyozov, 1961). At short times a nucleation barrier must be overcome before droplets which can grow form, and at later times the process leads to a power law growth of the characteristic length scale $L(t)$, i.e.

$$L(t) \propto t^{1/3} \tag{2.107}$$

for $d \geq 2$. Scaling behavior is also predicted for both the droplet size distribution $n_l(t)$ and the structure factor $S(q, t)$:

$$n_l(t) = (\bar{l}(t))^2 \tilde{n}(l/\bar{l}(t)), \qquad (l \to \infty, t \to \infty), \tag{2.108a}$$

$$S(t) = (L(t))^d \tilde{S}(qL(t)), \qquad (q \to \infty, t \to \infty), \tag{2.108b}$$

where $\bar{l} \propto t^{dx}$ is the mean cluster size and x is a characteristic exponent which is $1/3$ if conserved dynamics applies.

If, however, the initial quench is close to the critical point concentration, the state is unstable and the system evolves towards equilibrium by the formation of long wavelength fluctuations as shown in Fig. 2.12. The explicit shape of these structures will vary with model and with quench temperature; Fig. 2.12 is only intended to show 'typical' structures. The early stage of this process is called

spinodal decomposition and the late stage behavior is termed 'coarsening'. The linearized theory (Cahn and Hilliard, 1958; Cahn, 1961) predicts

$$S(q\,,t) = S(q\,,0)e^{2\omega(q)t} \tag{2.109}$$

where $\omega(q)$ is zero for the critical wavevector q_c. The linearized theory is invalid for systems with short range interactions but is approximately correct for systems with large, but finite, range coupling.

2.3.3 Critical slowing down at phase transitions

As a critical point T_c is approached the large spatial correlations which develop have long temporal correlations associated with them as well (van Hove, 1954). At T_c the characteristic time scales diverge in a manner which is determined in part by the nature of the conservation laws. This 'critical slowing down' has been observed in multiple physical systems by light scattering experiments (critical opalescence) as well as by neutron scattering. The seminal work by Halperin and Hohenberg (1977) provides the framework for the description of dynamic critical phenomena in which there are a number of different universality classes, some of which correspond to systems which only have relaxational behavior and some of which have 'true dynamics', i.e. those with equations of motion which are derived from the Hamiltonian. One consequence of this classification is that there may be different models which are in the same static universality class but which are in different dynamic classes. Simple examples include the Ising model with 'spin-flip' kinetics vs. the same model with 'spin-exchange' kinetics, and the Heisenberg model treated by Monte Carlo (stochastic) simulations vs. the same model solved by integrating coupled equations of motion. For relaxational models, such as the stochastic Ising model, the time-dependent behavior is described by a master equation

$$\partial P_n(t)/\partial t = -\sum_{n\neq m}[P_n(t)W_{n\to m} - P_m(t)W_{m\to n}], \tag{2.110}$$

where $P_n(t)$ is the probability of the system being in state 'n' at time t, and $W_{n\to m}$ is the transition rate for $n \to m$. The solution to the master equation is a sequence of states, but the time variable is a stochastic quantity which does not represent true time. A relaxation function $\phi(t)$ can be defined which describes time correlations *within* equilibrium

$$\phi_{MM}(t) = \frac{\langle M(0)M(t)\rangle - \langle M\rangle^2}{\langle M^2\rangle - \langle M\rangle^2}. \tag{2.111}$$

When normalized in this way, the relaxation function is 1 at $t = 0$ and decays to zero as $t \to \infty$. It is important to remember that for a system in equilibrium any time in the sequence of states may be chosen as the '$t = 0$' state. The asymptotic, long time behavior of the relaxation function is exponential, i.e.

$$\phi(t) \to e^{-t/\tau} \tag{2.112}$$

where the correlation time τ diverges as T_c is approached. This dynamic (relaxational) critical behavior can be expressed in terms of a power law as well,

$$\tau \propto \xi^z \propto \varepsilon^{-\nu z} \tag{2.113}$$

where ξ is the (divergent) correlation length, $\varepsilon = |1 - T/T_c|$, and z is the dynamic critical exponent. Estimates for z have been obtained for Ising models by epsilon-expansion RG theory (Bausch *et al.*, 1981) but the numerical estimates (Landau *et al.*, 1988; Wansleben and Landau, 1991; Ito, 1993) are still somewhat inconsistent and cannot yet be used with complete confidence.

A second relaxation time, the integrated relaxation time, is defined by the integral of the relaxation function

$$\tau_{\text{int}} = \int_0^\infty \phi(t)dt. \tag{2.114}$$

This quantity has particular importance for the determination of errors and is expected to diverge with the same dynamic exponent as the 'exponential' relaxation time.

One can also examine the approach to equilibrium by defining a non-linear relaxation function

$$\phi_M(t) = \frac{\langle M(t)\rangle - \langle M(\infty)\rangle}{\langle M(0)\rangle - \langle M(\infty)\rangle}. \tag{2.115}$$

The non-linear relaxation function also has an exponential decay at long times, and the characteristic relaxation time $\tau_{\text{nl}} = \int_0^\infty \phi_M(t)dt$ diverges with dynamic exponent z_{nl}. Fisher and Racz (1976) have shown, however, that there is only one independent exponent and that

$$z = z_{\text{nl}}^M + \beta/\nu, \tag{2.116}$$

or if the relaxation has been determined for the internal energy then

$$z = z_{\text{nl}}^E + (1 - \alpha)/\nu. \tag{2.117}$$

There are other systems, such as glasses and models with impurities, where the decay of the relaxation function is more complex. In these systems a 'stretched exponential' decay is observed

$$\phi \propto e^{-(t/\tau)^n}, \qquad n < 1 \tag{2.118}$$

and the behavior of τ may not be simple. In such cases, extremely long observation times may be needed to measure the relaxation time.

The properties of systems with true dynamics are governed by equations of motion and the time scale truly represents real time; since this behavior does not occur in Monte Carlo simulations it will not be discussed further at this point.

2.3.4 Transport coefficients

If some observable A is held constant and all 'flips' involve only local, e.g. nearest neighbor, changes, the Fourier components $A(q)$ can be described by a characteristic time

$$\tau_{AA}(q) = (D_{AA}q^2)^{-1} \tag{2.119}$$

where D_{AA} is a transport coefficient. In the simulation of a binary alloy, the concentrations of the constituents would be held fixed and D_{AA} would correspond to the concentration diffusivity. With different quantities held fixed, of course, different transport coefficients can be measured and we only offer the binary alloy model as an example. Equation (2.119) implies a very slow relaxation of long wavelength variations. Note that this 'hydrodynamic slowing down' is a very general consequence of the conservation of concentration and *not* due to any phase transition. If there is an unmixing critical point, see Fig. 2.11, then $D_{AA} \propto |\varepsilon|^{\gamma}$ and at T_c the relaxation time diverges as $\tau_{AA}(q) \propto q^{-(4-\eta)}$ (Hohenberg and Halperin, 1977).

2.3.5 Concluding comments: why bother about dynamics when doing Monte Carlo for statics?

Since importance sampling Monte Carlo methods correspond to a Markovian master equation by construction, the above remarks about dynamical behavior necessarily have some impact on simulations; indeed dynamical behavior can possibly affect the results for statics. For example, in the study of static critical behavior the critical slowing down will adversely affect the accuracy. We will return to this problem in Section 4.2.4.2, where we will also give an example. In the examination of hysteresis in the study of phase diagrams, etc. the long time scales associated with metastability are an essential feature of the observed behavior. Even if one simulates a fluid in the NVT ensemble away from any phase transition, there will be slow relaxation of long wavelength density fluctuations due to the conservation of density as in Eqn. (2.119). Thus, insight into the dynamical properties of simulations always helps to judge their validity.

REFERENCES

Barber, M. N., Pearson, R. B., Toussaint, D., and Richardson, J. L. (1985), *Phys. Rev. B* **32**, 1720.

Bausch, R., Dohm, V., Janssen, H. K., and Zia, R. K. P. (1981), *Phys. Rev. Lett.* **47**, 1837.

Becker, R. (1967), *Theory of Heat* (Springer, New York).

Binder, K. (1987), *Rep. Prog. Phys.* **50**, 783.

Cahn, J. W. (1961), *Acta Metall.* **9**, 795.

Cahn, J. W. and Hilliard, J. E. (1958), *J. Chem. Phys.* **28**, 258.

Callen, H. (1985), *Introduction to Thermodynamics and Thermostatics*, 2nd edn. (Wiley, New York).

Chen, K., Ferrenberg, A. M., and Landau, D. P. (1993), *Phys. Rev. B* **48**, 239 and references therein.
Coddington, P. D. (1994), *Int. J. Mod. Phys. C* **5**, 547.
Domany, E., Schick, M., Walker, J. S., and Griffiths, R. B. (1980), *Phys. Rev. B* **18**, 2209.
Dunkel, J. and Hilbert, S. (2006), *Physica A* **370**, 390.
Feller, W. (1968), *An Introduction to Probability Theory and its Applications*, vol. 1 (J. Wiley and Sons, New York).
Ferrenberg, A. M., Landau, D. P., and Binder, K. (1991), *J. Stat. Phys.* **63**, 867.
Fisher, M. E. (1974), *Rev. Mod. Phys.* **46**, 597.
Fisher, M. E. and Racz, Z. (1976), *Phys. Rev. B* **13**, 5039.
Gentle, J. E. (2003), *Random Number Generation and Monte Carlo Methods*, 2nd edn. (Springer, Berlin).
Griffiths, R. B. (1970), *Phys. Rev. Lett.* **24**, 715.
Gunton, J. D., San Miguel, M., and Sahni, P. S. (1983), in *Phase Transitions and Critical Phenomena*, vol. 8, eds. C. Domb and J. L. Lebowitz (Academic Press, London), p. 267.
Hohenberg, P. and Halperin, B. (1977), *Rev. Mod. Phys.* **49**, 435.
Ito, N. (1993), *Physica A* **196**, 591.
James, F. (1994), *Comput. Phys. Commun.* **79**, 111.
Kadanoff, L. P., Goetze, W., Hamblen, D., Hecht, R., Lewis, E. A. S., Palciauskas, V. V., Rayl, M., Swift, J., Aspenes, D., and Kane, J. (1967), *Rev. Mod. Phys.* **39**, 395.
Kalos, M. H. and Whitlock, P. A. (1986), *Monte Carlo Methods*, vol. 1 (Wiley and Sons, New York).
Katzgraber, H. G. (2011), in *Modern Computational Science 11*, eds. R. Leidl and A. K. Hartmann (BIS Verlag, Oldenburg), p. 73.
Kirkpatrick, S. and Stoll, E. (1981), *J. Comput. Phys.* **40**, 517.
Knuth, D. (1969), *The Art of Computer Programming*, vol. 2 (Addison-Wesley, Reading, MA).
Kornblit, A. and Ahlers, G. (1973), *Phys. Rev. B* **8**, 5163.
Kosterlitz, J. M. and Thouless, D. J. (1973), *J. Phys. C* **6**, 1181.
L'Ecuyer, P., and Simard, R. (2007), *ACM Trans. Math. Softw.* **33**, 4.
Landau, D. P. (1983), *Phys. Rev. B* **27**, 5604.
Landau, D. P. (1994), *J. Appl. Phys.* **73**, 6091.
Landau, D. P. and Swendsen, R. H. (1981), *Phys. Rev. Lett.* **46**, 1437.
Landau, D. P., Tang, S., and Wansleben, S. (1988), *J. de Physique* **49**, C8–1525.
Landau, L. D. and Lifshitz, E. M. (1980), *Statistical Physics*, 3rd edn. (Pergamon Press, Oxford).
Lawrie, I. D. and Sarbach, S. (1984), in *Phase Transitions and Critical Phenomena*, vol. 9, eds. C. Domb and L. J. Lebowitz (Academic Press, New York), p. 1.
Lifshitz, I. M. and Slyozov, V. V. (1961), *J. Chem. Solids* **19**, 35.
Lüscher, M. (1994), *Comput. Phys. Commun.* **79**, 100.
Manssen, M., Weigel, M., and Hartmann, A. K. (2012), *Eur. Phys. J. Special Topics* **210**, 53.
Marsaglia, G. (1968), *Proc. Natl. Acad. Sci.* **61**, 25.
Marsaglia, G. (1985), in *Computer Science and Statistics: The Interface*, ed. L. Billard (Elsevier, Amsterdam).
Marsaglia, G. (unpublished), *Diehard Battery of Tests for Random Number Generators*, http://www.stat.fsu.edu/pub/diehard/.
Marsaglia, G. and Zaman, A. (1991), *Ann. Appl. Prob.* **1**, 462.
Marsaglia, G., Narasimhan, B., and Zaman, A. (1990), *Comput. Phys. Comm.* **60**, 345.
Nelson, D. R., Kosterlitz, J. M., and Fisher, M. E. (1974), *Phys. Rev. Lett.* **33**, 813.

Nienhuis, B. (1982), *J. Phys. A: Math. Gen.* **15**, 199.
Onsager, L. (1944), *Phys. Rev.* **65**, 117.
Pelissetto, A. and Vicari, E. (2002), *Physics Reports* **368** (6), 549.
Potts, R. B. (1952), *Proc. Camb. Philos. Soc.* **48**, 106.
Privman, V., Hohenberg, P. C., and Aharony, A. (1991), in *Phase Transitions and Critical Phenomena*, vol. 14, eds. C. Domb and J. L. Lebowitz (Academic Press, London).
Riedel, E. K. and Wegner, F. J. (1972), *Phys. Rev. Lett.* **29**, 349.
Sachdev, S. (1999), *Quantum Phase Transitions* (Cambridge University Press, Cambridge).
Stanley, H. E. (1971), *An Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford).
Stryjewski, E. and Giordano, N. (1977), *Adv. Physics* **26**, 487.
van Hove, L. (1954), *Phys. Rev.* **93**, 1374.
von Neumann, J. (1951), *J. Res. Nat. Bur. Stand.* **12**, 36.
Wansleben, S. and Landau, D. P. (1991), *Phys. Rev. B* **43**, 6006.
Wegner, F. J. and Riedel, E. K. (1973), *Phys. Rev. B* **7**, 248.
Wilson, K. G. (1971), *Phys. Rev. B* **4**, 3174, 3184.
Wu, F. Y. (1982), *Rev. Mod. Phys.* **54**, 235.
Yeomans, J. (1992), *Statistical Mechanics of Phase Transitions* (Oxford University Press, Oxford).
Zinn-Justin, J. and LeGuillou, J.-C. (1980), *Phys. Rev. B* **21**, 3976.