5
Applications to Statistical Mechanics systems

5.1
Introduction

One of the most important applications of the Monte Carlo sampling techniques appears in the field of Statistical Mechanics, which deals with systems with a large number of degrees of freedom N. A system is described by a set of coordinates $X = (x_1, \ldots, x_N)$. These variables x_i are typically positions, or angles, but more complicated examples of generalized coordinates exist in the literature. Each coordinate x_i has an associated conjugate momentum p_i and the whole set of X and $P = (p_1, \ldots, p_N)$ variables is required to fully specify the state of the system. We will denote by $\Gamma = (X, P)$ the combined set of $2N$ coordinate and momentum variables and a point in the Γ "phase space" is a "microscopic configuration". It is essential to remember that the simplest macroscopic system will be described by an enormously large number of variables N. A measure of how large this number can be in typical situations is Avogadro's number $N_A = 6.022 \times 10^{23}$, the number of molecules in a mol, usually a few grams, of substance.

A very important function is the Hamiltonian $\mathcal{H}(X, P) = \mathcal{H}(x_1, \ldots, x_N, p_1, \ldots, p_N)$. It is important in many senses. First of all, it determines, via Hamilton's equations

$$\frac{dx_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i}, \tag{5.1}$$

$$\frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial x_i}, \qquad i = 1, \ldots, N, \tag{5.2}$$

the time evolution $x_i(t), p_i(t), i = 1, \ldots, N$. All we need to do is to solve this set of $2N$ differential equations given some initial condition at, say, time $t = 0$. Of course, a tremendous (and impossible) task in most cases.

Second, and more importantly to us, is that the Hamiltonian can also be used to determine the probability density function of observing at thermal equilibrium some set of values for the coordinates and momenta $\Gamma = (X, P)$, i.e. the probability of a microscopic configuration. Boltzmann and Gibbs were the ones to show that such a pdf is given by what is nowadays called the Boltzmann factor, $e^{-\beta \mathcal{H}}$, in the following manner

$$f(\Gamma) = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}}, \tag{5.3}$$

where

$$\mathcal{Z} = \int d\Gamma \, e^{-\beta \mathcal{H}} \tag{5.4}$$

is the normalization factor and $\beta = 1/kT$ is the inverse of the temperature rescaled by Boltzmann's constant k. Just not to hide anything, we can write this normalization factor in full

$$\mathcal{Z} = \int dx_1 \dots dx_N \, dp_1 \dots dp_N \, e^{-\beta \mathcal{H}(x_1,\dots,x_N,p_1,\dots,p_N)}. \tag{5.5}$$

In fact, this normalization factor by itself is very important, so important that it has a name, "the partition function". The reason of its importance is that the probability of a configuration is something difficult to determine experimentally, while typical measures concern the so-called thermodynamic potentials: Helmholtz free energy F, internal energy $\mathcal{U}$, enthalpy H, or the equation of state giving the pressure P as as function of volume V and temperature, or the specific heat at constant volume C_V, or the isothermal compressibility κ_T, etc. The framework of Statistical Mechanics tells us that all these quantities can be derived from the only knowledge of the partition function as a function of the volume V, the number of particles $\mathbf{N}$[1)] and the temperature T. For example, Helmholtz's free energy is given by

$$F(\mathbf{N}, V, T) = -kT \log \mathcal{Z}, \tag{5.6}$$

the internal energy is

$$\mathcal{U} = \left(\frac{\partial F/T}{\partial (1/T)} \right)_{V,\mathbf{N}}, \tag{5.7}$$

and the specific heat at constant volume,

$$C_V = \left(\frac{\partial \mathcal{U}}{\partial T} \right)_{V,\mathbf{N}}. \tag{5.8}$$

The entropy can be computed from $S = (\mathcal{U} - F)/T$ or directly:

$$S = -\left(\frac{\partial F}{\partial T} \right)_{V,\mathbf{N}} \tag{5.9}$$

and so on. This "recipe" of Statistical Mechanics is extremely difficult to carry on in practice as the integrals involved in the definition of the partition function (5.5) can only be performed for a limited number of simple examples: the gas of non-interacting particles, a system of independent harmonic oscillators, etc. and a limited number of not so simple examples: Gaussian free model and the Ising model for ferromagnetism being the most noticeable ones. Furthermore, the framework of Statistical Mechanics shows that some interesting macroscopic observable phenomena only occur in the limit of the number of degrees of freedom N tending to infinite, which makes the calculation of the partition function usually even harder.

1) The number of particles $\mathbf{N}$ is not the same as the number of degrees of freedom N, as each particle can have, in general, more than one degree of freedom.

An alternative approach is to work not with the partition function directly, but with the pdf of the configurations, $f(\Gamma)$. In this approach, some observables are defined in terms of averages with respect to this pdf. The average of any function of the microscopic variables $G(\Gamma)$ is defined in the usual way:

$$\langle G \rangle = \int d\Gamma G(\Gamma) f(\Gamma). \tag{5.10}$$

There are many examples. The internal energy can be computed as the average value of the Hamiltonian,

$$\mathcal{U} = \langle \mathcal{H} \rangle. \tag{5.11}$$

The specific heat, $C_V = \left(\frac{\partial \mathcal{U}}{\partial T}\right)_{\mathbf{N},V}$, can also be obtained using higher-order moments of the Hamiltonian:

$$C_V = k\beta^2 \left[\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2\right] \tag{5.12}$$

and so on. Unfortunately, not all relevant magnitudes can be computed this way. For instance, the entropy S can not be interpreted as the average value of some known function G[2)].

Many problems of Statistical Mechanics in equilibrium can then be reduced to the calculation of averages using the pdf (5.3). This probabilistic description (which lies at the core of Statistical Mechanics) is, in some cases, independent on whether the dynamical variables themselves satisfy or not Hamilton's equations. For instance, in problems of magnetism, a very successful approach consists in considering that the variables x_i represent microscopic magnetic moments which interact amongst themselves. In the simplest version, known as Ising model, these microscopic variables can take only two possible values $x_i \equiv \mu s_i$, being μ the unit of magnetic moment and $s_i = \pm 1$ a rescaled variable (the Ising or spin variable). There are no momenta-like variables p_i associated to these magnetic moment variables and there are no Hamilton's equations. It does not even make sense to compute the time derivative of a non-continuous variable s_i that can only take two possible values. Still, the microscopic variables s_i interact via the so-called "Hamiltonian" function $\mathcal{H}(s_1, \ldots, s_N)$ and the probability of observing a particular configuration $S = (s_1, \ldots, s_N)$ is $f(S) = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}}$. The partition function now is not the integral over all values of s_i but, as s_i can take only two values $s_i = \pm 1$, it is computed as a sum:

$$\mathcal{Z} = \sum_{s_1 = \pm 1} \cdots \sum_{s_N = \pm 1} e^{-\beta \mathcal{H}}. \tag{5.13}$$

The Hamiltonian $\mathcal{H}$ takes into account the magnetic interactions, and it is typically simplified in order to consider only the interactions that occur between these Ising

2) It is indeed possible to derive the formula $S = -\int d\Gamma f(\Gamma) \log[f(\Gamma)] + C$, being C a constant. Though formally this could be thought of as the average $S = -\langle \log f \rangle + C$, the truth is that in order to perform this average, we must know the pdf $f(\Gamma)$, *including the normalization constant* $\mathcal{Z}$, which is usually impossible. If we knew the partition function, we would not need any further integrals in order to compute the entropy, we'd simply use (5.6)-(5.9).

variables s_i which are close in space, neglecting the magnetic interactions with variables which are farther apart than some minimum cut-off distance. We will see specific examples later.

Once we have reduced the problem of equilibrium Statistical Mechanics to the calculation of averages, it is clear which should be an efficient numerical approach: replace the true average (5.10) by a sample average

$$\langle G\rangle = \mu[G] \pm \frac{\sigma[G]}{\sqrt{M}}(2\tau_G + 1), \tag{5.14}$$

with

$$\mu[G] = \frac{1}{M}\sum_{k=1}^{M} G(\Gamma_k), \tag{5.15}$$

$$\sigma^2[G] = \frac{1}{M}\sum_{k=1}^{M} G(\Gamma_k)^2 - (\mu[G])^2, \tag{5.16}$$

being Γ_k, $k = 1, \ldots, M$ the set of generated configurations and τ_G the associated autocorrelation time of G obtained from the autocorrelation function $\rho_G(i)$. The key point is now the generation of characteristic configurations Γ_k distributed according to $f(\Gamma)$ in a problem with many variables. But this is precisely what we have claimed that Monte Carlo algorithms are good at! We now rephrase the dynamical methods of last chapter within the framework of Statistical Mechanics problems.

In the dynamical methods we proposed a change from configuration Γ to configuration Γ' taken from a pdf $g(\Gamma'|\Gamma)$. This proposal was then accepted with probability $h(\Gamma'|\Gamma)$. In order to ensure that the stationary distribution is $f(\Gamma)$, it is sufficient to demand that these conditional functions satisfy the detailed balance condition:

$$g(\Gamma'|\Gamma)h(\Gamma'|\Gamma)f(\Gamma) = g(\Gamma|\Gamma')h(\Gamma|\Gamma')f(\Gamma'), \tag{5.17}$$

with now $f(\Gamma)$ given by (5.3). A solution to this functional equation is given by the Metropolis algorithm, in which a proposal probability $g(\Gamma'|\Gamma)$ is first selected and then the acceptance probability is

$$h(\Gamma'|\Gamma) = \min\left[1, q(\Gamma'|\Gamma)\right], \tag{5.18}$$

where

$$q(\Gamma'|\Gamma) = \frac{g(\Gamma|\Gamma')f(\Gamma')}{g(\Gamma'|\Gamma)f(\Gamma)}. \tag{5.19}$$

Most algorithms[3)] assume that $g(\Gamma'|\Gamma)$ is a symmetric function, $g(\Gamma'|\Gamma) = g(\Gamma|\Gamma')$. In this case, and after replacing the expression for $f(\Gamma)$ we obtain

$$q(\Gamma'|\Gamma) = \frac{\mathcal{Z}^{-1}e^{-\beta\mathcal{H}(\Gamma')}}{\mathcal{Z}^{-1}e^{-\beta\mathcal{H}(\Gamma)}} = e^{-\beta\Delta\mathcal{H}}, \tag{5.20}$$

3) But not all. A notable exception is the Hybrid Monte Carlo algorithm explained in chapter 10.

with $\Delta\mathcal{H} = \mathcal{H}(\Gamma') - \mathcal{H}(\Gamma)$ the change in energy involved in the proposal $\Gamma \to \Gamma'$. Note that the partition function $\mathcal{Z}$ disappears from the expression for $q(\Gamma'|\Gamma)$. The acceptance probability becomes,

$$h(\Gamma'|\Gamma) = \min(1, e^{-\beta\Delta\mathcal{H}}). \tag{5.21}$$

This is the original proposal of the celebrated paper by Metropolis *et al*[7]. It has an intuitive physical interpretation. The characteristic configurations at equilibrium are those that minimize Helmholtz's free energy $F = \mathcal{U} - TS$, a balance between the internal energy $\mathcal{U}$ (which tends to a minimum) and entropy S (which tends to a maximum). As $\Delta\mathcal{H} \leq 0$ implies $h(\Gamma'|\Gamma) = 1$, this balance is achieved by (i) accepting all proposals $\Gamma \to \Gamma'$ in which the energy is reduced, and (ii) accepting those proposals in which energy increases $\Delta\mathcal{H} > 0$ with a probability $e^{-\beta\Delta\mathcal{H}}$. As $\beta = 1/kT$, when $T \to 0$ the probability of accepting a proposal that increases the energy tends to 0. On the contrary, when $T \to \infty$ it is $\beta \to 0$, the acceptance probability tends to 1, and every proposal is accepted independently on the energy cost.

Another solution to the detailed balance condition is that of Glauber

$$h(\Gamma'|\Gamma) = \frac{q(\Gamma'|\Gamma)}{1 + q(\Gamma'|\Gamma)}, \tag{5.22}$$

or, using (5.20),

$$h(\Gamma'|\Gamma) = \frac{1}{1 + e^{\beta\Delta\mathcal{H}}}, \tag{5.23}$$

but other choices for the acceptance probability h are still possible, as discussed in section 4.4.

5.2 Average acceptance probability

We now rewrite result (4.91) in order to derive the average acceptance probability in the case that the pdf is the exponential of a Hamiltonian: $f_{\hat{\mathbf{x}}}(\Gamma) = \mathcal{Z}^{-1} e^{-\beta\mathcal{H}(\Gamma)}$, and assuming a symmetric proposal $g(\Gamma|\Gamma') = g(\Gamma'|\Gamma)$:

$$\left\langle e^{-\beta\Delta\mathcal{H}} \right\rangle_{\text{st}} = 1, \tag{5.24}$$

being $\Delta\mathcal{H} = \mathcal{H}(\Gamma') - \mathcal{H}(\Gamma)$ the change of energy involved in the proposal $\Gamma \to \Gamma'$. If we use now Jensen's inequality $\langle e^{-z} \rangle \geq e^{-\langle z \rangle}$, valid for any random variable z, with $z = \beta\Delta\mathcal{H}$, we derive that in the steady state the average value of the proposed changes of energy is always greater than zero, $\langle \Delta\mathcal{H} \rangle > 0$.

Although it is possible to be more general, we present here a simplified treatment which assumes that the distribution of proposed changes of energy $z = \beta\Delta\mathcal{H}$ can be well approximated by a Gaussian distribution of average μ and variance σ^2. As it is known that, for a Gaussian distribution, it is $\langle e^{-z} \rangle = e^{-\mu + \sigma^2/2}$ and we have

proved that $\langle e^{-z}\rangle = 1$, we derive that this Gaussian assumption is consistent if we take $\sigma^2 = 2\mu$. Let us now compute the average value of the Metropolis acceptance probability $h(\Gamma'|\Gamma) = \min[1, e^{-\beta\Delta\mathcal{H}}]$. Under the assumption that $z = \beta\Delta\mathcal{H}$ follows a Gaussian distribution we obtain for the average value:

$$\langle h(\Gamma'|\Gamma)\rangle_{\rm st} = \int_{-\infty}^{\infty} dz \min[1, e^{-z}] \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}. \tag{5.25}$$

Performing the integral and replacing $\sigma^2 = 2\mu$ we obtain the simple result $\langle h(\Gamma'|\Gamma)\rangle_{\rm st} = \text{erfc}\left(\frac{\sqrt{\mu}}{2}\right)$ or

$$\langle h(\Gamma'|\Gamma)\rangle_{\rm st} = \text{erfc}\left(\frac{\sqrt{\beta\langle\Delta\mathcal{H}\rangle}}{2}\right), \tag{5.26}$$

being $\text{erfc}(z) = 1 - \text{erf}(z)$ the complementary error function. It is worth noting that in the limit of large $\langle\Delta\mathcal{H}\rangle$, we can use $\text{erfc}(z) \underset{z\to\infty}{\longrightarrow} \frac{e^{-z^2}}{z\sqrt{\pi}}$ [4] to obtain the asymptotic result:

$$\langle h(\Gamma'|\Gamma)\rangle_{\rm st} \longrightarrow \frac{2e^{-\frac{\beta\langle\Delta\mathcal{H}\rangle}{4}}}{\sqrt{\pi\beta\langle\Delta\mathcal{H}\rangle}}. \tag{5.27}$$

valid for $\Delta\mathcal{H} \to \infty$ and showing that the average acceptance probability $\langle h(\Gamma'|\Gamma)\rangle$ goes to zero *exponentially* with the average (positive) energy change $\langle\Delta\mathcal{H}\rangle$. As $\mathcal{H}$ is usually an extensive quantity proportional to the number of degrees of freedom N, it is important to devise proposals in which the change of energy $\Delta\mathcal{H}$ can be kept small, of $O(1)$ instead of $O(N)$, otherwise the small average acceptance probability will yield a very large correlation time and the corresponding large statistical errors will make the estimator useless. The standard trick to keep a reasonable acceptance probability is to propose new configurations Γ' in which only a few degrees of freedom have been modified from configuration Γ. Collective updates in which all degrees of freedom change at once but still the acceptance probability is non vanishingly small are much more difficult to devise. We will devote chapter 10 and appendix D to explain some collective update algorithms, and restrict ourselves in the rest of the chapter to algorithms in which the proposal involves the change of a few (maybe only one) variable.

Let us know explain some applications of the Monte Carlo algorithms to the study of particular systems of interest in Statistical Mechanics.

5.3 Interacting particles

The first case we will consider is that of interacting particles without internal degrees of freedom. This means that all we need to specify the state are the spatial locations

4) The limiting expression has an error smaller than 10^{-6} for $z > 3$.

$\vec{r}_i$ of the $i = 1, \ldots, \mathbf{N}$ particles, as well as the associated momenta $\vec{p}_i = m\vec{v}_i$, being m the mass of the particle (assumed identical) and $\vec{v}_i$, the velocity. In general $\vec{r}_i = (x_i, y_i, z_i)$ and $\vec{p}_i = (p_i^x, p_i^y, p_i^z)$ are three-dimensional vectors, although other spatial dimensions can be considered in specific cases. We can think of the particles as perfect spheres (or rods in one dimension and disks in two dimensions) moving around while interacting with other particles.

The Hamiltonian consists of two terms, corresponding to the kinetic, $\mathcal{T}$, and potential, $\mathcal{V}$, energy. In the potential energy, we must consider interactions between all possible pairs of particles (i, j), which, for convention, we order using $i < j$. As the particles have no internal degrees of freedom, the potential interaction $v(\vec{r}_i, \vec{r}_j)$ between particles i and j depends only on the locations of the particle (we assume it does not depend on their velocities). The Hamiltonian is, then,

$$\mathcal{H}(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}}; \vec{p}_1, \ldots, \vec{p}_{\mathbf{N}}) = \mathcal{T}(\vec{p}_1, \ldots, \vec{p}_{\mathbf{N}}) + \mathcal{V}(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}}), \tag{5.28}$$

with

$$\mathcal{T} = \sum_{i=1}^{\mathbf{N}} \frac{\vec{p}_i{}^2}{2m}, \tag{5.29}$$

$$\mathcal{V} = \sum_{i<j} v(\vec{r}_i, \vec{r}_j). \tag{5.30}$$

We see that the pdf $e^{-\beta\mathcal{H}}$ can be split naturally as:

$$e^{-\beta\mathcal{H}} = \left[\prod_{i=1}^{\mathbf{N}} e^{-\beta \frac{\vec{p}_i{}^2}{2m}}\right] \times e^{-\beta \sum_{i<j} v(\vec{r}_i, \vec{r}_j)} \tag{5.31}$$

$$= \left[\prod_{i=1}^{\mathbf{N}} e^{-\frac{(p_i^x)^2}{2m/\beta}} e^{-\frac{(p_i^y)^2}{2m/\beta}} e^{-\frac{(p_i^z)^2}{2m/\beta}}\right] \times e^{-\beta \sum_{i<j} v(\vec{r}_i, \vec{r}_j)} \tag{5.32}$$

which indicates that, from the statistical point of view, each one of the momenta coordinates (p_i^x, p_i^y, p_i^z) are independently distributed with a Gaussian distribution of zero mean and variance $m/\beta = kTm$. This independence allows us to obtain analytically some averages of interest. For instance, the average value of the kinetic energy is

$$\left\langle \sum_{i=1}^{\mathbf{N}} \frac{\vec{p}_i{}^2}{2m} \right\rangle = 3\mathbf{N}kT, \tag{5.33}$$

as each of the $3\mathbf{N}$ Gaussian variables contributes a factor kTm (energy equipartition theorem). On the other hand, averages of functions $G(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}})$ which depend on the coordinates are much more difficult to perform analytically due to the interaction terms. Here is where the numerical methods are useful. The formal expression is

$$\langle G(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}}) \rangle = C^{-1} \int d\vec{r}_1 \ldots d\vec{r}_{\mathbf{N}} G(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}}) e^{-\beta\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}})}, \tag{5.34}$$

or an average with respect to the pdf

$$f(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}}) = C^{-1} e^{-\beta\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_{\mathbf{N}})}, \tag{5.35}$$

being $C = \int d\vec{r}_1 \dots d\vec{r}_{\mathbf{N}} e^{-\beta \mathcal{V}(\vec{r}_1,\dots,\vec{r}_{\mathbf{N}})}$, the normalization constant.

To perform these averages numerically, we generate configurations of positions $X \equiv (\vec{r}_1, \dots, \vec{r}_{\mathbf{N}})$ distributed according to this pdf and approximate the average (5.34) by the sample average, including an estimation of the error. In order to generate the configurations, we can use, for example, a Metropolis algorithm, in which we propose a change from $X \to X'$ according to some distribution $g(X'|X)$. We have much freedom in choosing the new configuration X', but we must do so such that the resulting acceptance probability $h(X'|X)$ is not vanishingly small. As explained before, we choose a proposal X' that differs from X in just a few variables. For instance, we can choose, most naturally, to change the position $\vec{r}_i \to \vec{r}_i'$ of a single particle, randomly chosen amongst the $\mathbf{N}$ particles. In this way, and similarly to (4.71), $g(X'|X)$ is constructed from $g(\vec{r}_i'|\vec{r}_i)$ which depends only on the coordinates of the randomly chosen particle i. We can choose, for instance (but we stress that we have a lot of freedom in this proposal step), to change every cartesian coordinate $(x_i, y_i, z_i) \to (x_i', y_i', z_i')$ such that x_i', y_i' and z_i' are drawn from a uniform distribution in the intervals $(x_i - \Delta, x_i + \Delta)$, $(y_i - \Delta, y_i + \Delta)$ and $(z_i - \Delta, z_i + \Delta)$, respectively. As this proposal is symmetrical $g(X'|X) = g(X|X')$, the detailed balance condition is

$$h(X'|X) e^{-\beta \mathcal{V}(X)} = h(X|X') e^{-\beta \mathcal{V}(X')}, \tag{5.36}$$

in terms of the potential energy $\mathcal{V}$ only. We can take, for instance, Metropolis solution

$$h(X'|X) = \min\left[1, e^{-\beta \Delta \mathcal{V}}\right] \tag{5.37}$$

being $\Delta \mathcal{V} = \mathcal{V}(X') - \mathcal{V}(X)$, the change in the potential energy induced by the proposed change $X \to X'$. As only the position $\vec{r}_i$ is modified, the change is:

$$\Delta \mathcal{V} = \sum_{j \neq i} \left[v(\vec{r}_i', \vec{r}_j) - v(\vec{r}_i, \vec{r}_j) \right]. \tag{5.38}$$

Usually, the potential interaction depends on the separation distance $r_{ij} = |\vec{r}_i - \vec{r}_j|$ between particle i and j, $v(\vec{r}_i, \vec{r}_j) = v(r_{ij})$. This is the case, for example, of the celebrated Lennard-Jones potential

$$v(r) = v_0 \left[\left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^{6} \right], \tag{5.39}$$

being v_0 and r_0 parameters of the potential.

Note that, in this case, and for a large number of particles $\mathbf{N}$, the calculation of the sum (5.38) can be very expensive from the computational point of view as a sum of $\mathbf{N}$ terms must be calculated every time a single position is proposed to change. This is another reason why collective updates proposals, in which all variables $(\vec{r}_1, \dots, \vec{r}_{\mathbf{N}})$ are proposed to change simultaneously are more effective. A particular type of collective updating, suitable for this kind of systems and known as Hybrid Monte Carlo, will be explained in chapter 10.

In some cases, the potential interaction $v(\vec{r}_i, \vec{r}_j)$ is such that it vanishes whenever the separation distance $|\vec{r}_i - \vec{r}_j|$ is larger than a cut-off R. This is typically implemented in the Lennard-Jones potential as:

$$v(r) = \begin{cases} v_0 \left[\left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^{6} \right], & r \leq R, \\ 0, & r > R. \end{cases} \tag{5.40}$$

In this case, the sum (5.38) has a much smaller number of terms,

$$\Delta\mathcal{V} = \sum_{j, |\vec{r}\,'_i - \vec{r}_j| < R, |\vec{r}_i - \vec{r}_j| < R} \left[v(\vec{r}\,'_i, \vec{r}_j) - v(\vec{r}_i, \vec{r}_j) \right]. \tag{5.41}$$

However, some non-trivial bookkeeping is necessary in order to keep track of which particles j are at a distance less than the cutoff from the particle i whose position is proposed to change to $\vec{r}\,'_i$.

The main features of the Lennard-Jones potential are the presence of a repulsion term (r^{-12}) at short distances and an attraction term (r^{-6}) decreasing to zero at long distances. These features can be captured by simpler models. For instance, the potential

$$v(r) = \begin{cases} \infty, & r \leq 2\sigma, \\ -v_0, & 2\sigma < r \leq R, \\ 0, & r > R, \end{cases} \tag{5.42}$$

signifying that the attraction between particles only occurs if they are closer than a distance R and that the repulsion energy is infinite if particles try to get closer than a distance 2σ, modeling the so-called *hard-core* repulsion. This can be imagined as each particle being a sphere of radius σ, such that the energetic cost of bringing two particles closer than a distance of one diameter, 2σ, is infinite. When only the repulsion term is taken into consideration, i.e. when the potential is:

$$v(r) = \begin{cases} \infty, & r \leq 2\sigma, \\ 0, & r > 2\sigma, \end{cases} \tag{5.43}$$

we talk about a system of hard-spheres. In this particular case, the energy change $\Delta\mathcal{V}$ is either zero, if the new position $\vec{r}\,'_i$ is such that the particle does not overlap with any other, or infinite, in case there is one particle j with which it overlaps. The acceptance probability is, hence,

$$h(X'|X) = \min\left[1, e^{-\beta\Delta\mathcal{V}}\right] = \begin{cases} 0 & \text{if exists } j \text{ such that } |\vec{r}\,'_i - \vec{r}_j| < 2\sigma, \\ 1, & \text{otherwise.} \end{cases} \tag{5.44}$$

In other words, movements are accepted if and only if they do not lead to overlaps between particles. Again, the program can be make much more efficient by making an *a priori* list containing the particles with which particle i can overlap after it is moved. Since the maximum change in the modulus of the position is $\sqrt{3}\Delta$, the list must include all particles which are at a distance shorter than $2\sigma + \sqrt{3}\Delta$ from particle

i. Simple as it might look, the system of hard spheres has been extensively studied both from the theoretical and numerical points of view. Very efficient codes have been developed to this end, and we refer the interested reader to the more specialized literature in this subject.

5.4 Ising model

We have already introduced the Ising model. Although it has been used in very different contexts (from phase separation in binary metal alloys to segregation in urban communities), its most direct application (and the one it was originally introduced for) is that of magnetic materials. Imagine a substance displaying the para-ferromagnetic transition. This is to say, one magnet that at high temperatures (above the so-called Curie temperature) loses its spontaneous magnetization[5)]. We can simplify the complicated structure of the magnetic solid by a perfect, regular lattice and assume that in every site i of this lattice lies a microscopic magnet, capable of taking two possible values $x_i = \mu s_i = \pm\mu$ for the magnetic moment. Here s_i is called the "spin"-variable (a name reflecting the fact that the magnetism has its origin in the individual atomic spins). The value $s_i = +1$ means that the magnet points upwards in an arbitrary, Z, direction, while a value $s_i = -\mu$ indicating that the magnet points downwards in the opposite direction. A microscopic configuration $S = (s_1, \ldots, s_N)$ is a set of values for the N spin variables.

The last ingredient is an energy function, a Hamiltonian[6)] reflecting the magnetic interaction between the spins. In the ferromagnetic materials, the interaction is such that it favors two spins to point in the same direction. This is reflected in a potential interaction between spins at sites i and j equal to $-J_{ij}s_is_j$, with J_{ij}, the *coupling constant* between sites i and j, a positive parameter. Hence, if s_i and s_j are both parallel, i.e. both take the value $+1$ or the value -1, then the interaction energy is $-J_{ij}$, while if they are antiparallel, one of them $+1$ and the other -1, the interaction energy is $+J_{ij}$, higher (for $J_{ij} > 0$) than in the case of parallel alignments. The interaction energy J_{ij} decays with the distance between sites i and j. An important simplification of the model is to assume that the interaction is very short ranged and only occurs for those spins which are sufficiently close in the regular lattice[7)]. What is meant by "sufficiently close" depends on the type of lattice, but usually one adopts the point of view that only those spins which lie apart the minimum distance dictated by the lattice are capable to interact between them. Those spins are then said to be "nearest neighbors" in the lattice. In figure 5.1 we plot some common lattices and the underlying structure of nearest neighbors.

5) We might not be very familiar with this situation, as the most common magnet, iron, loses its magnetization at around 770 C, certainly not an everyday temperature

6) We stress again that there are no Hamilton's equations associated to this function.

7) Another simplification of the model goes in the opposite direction and assumes a fully-connected lattice in which all spins interact with all others with the same energy. This is the so-called mean-field version of the Ising model. It is less realistic, but its main advantage is that it can be solved analytically.

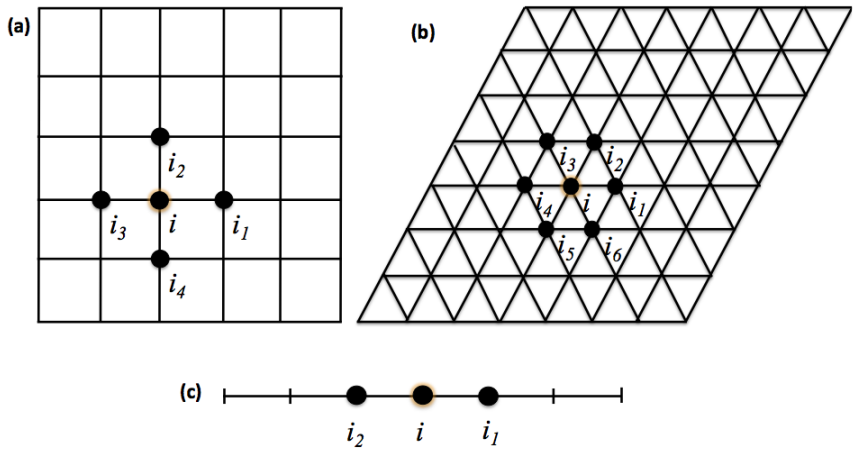

Figure 5.1 In the square-lattice, panel (a), node i has four nearest-neighbors: i_1, i_2, i_3, i_4; in the triangular lattice, panel (b), node i has six nearest-neighbors: i_1, i_2, i_3, i_4, i_5, i_6; in the linear chain, panel (c), node i has two nearest-neighbors: i_1, i_2.

The Hamiltonian, finally, takes into account the possible existence of a magnetic field H with which spins tend to align. This is reflected by a term $-Hs_i$ which takes its minimum value whenever the sign of H and s_i coincide. With all these considerations in mind, the Hamiltonian of the Ising model is:

$$\mathcal{H}(s_1, \ldots, s_N) = -J \sum_{\langle i,j \rangle} s_i s_j - H \sum_i s_i, \tag{5.45}$$

where the notation $\langle i, j \rangle$ indicates precisely all the pairs of sites i, j which are nearest neighbors in the chosen network.

In the absence of a magnetic field, $H = 0$, The basic phenomenology of the Ising model is that the tendency to align parallel dominates at low temperatures and there is a vast majority of spins pointing in the same direction, a situation identified with macroscopic order. Whether this direction is up, $s_i = +1$, or down, $s_i = -1$, depends on many things (for example, the initial conditions). The choosing of one of the two, otherwise equivalent, directions is an example of "symmetry breaking". At temperatures above a critical value T_c, the disordering role dominates and approximately half of the spins point upwards and half downwards. The symmetry has been restored. This competition between an ordering agent (the coupling constant) and a disordering one (the temperature) and the resulting transition between order and disorder is arguably the simplest example of a "phase transition", the phases being the ordered state at low temperature (the ferromagnetic phase) and the disordered state at high temperatures (the paramagnetic phase). This competition between ordering and disordering agents and the resulting phase transition phenomenology can be found in many situations, not just of physical interest, and the Ising model is then used

as a paradigmatic case study. An example far from the topic of magnetic materials occurs in the field of opinion formation, where positive and negative opinions over a given topic can coexist with approximately half of the population supporting one or the other, or one of them can dominate. Whatever the interpretation, it is usual to keep the magnetic notation and talk about a "paramagnetic" and a "ferromagnetic" phase, or "spin interaction", even if the interpretation of the model is far from the original domain of magnetism for which it was designed.

Which averages of functions G of the spin variables $(s_1, \ldots, s_N)$ are useful to compute in the Ising model? Without doubt the most important one is the magnetization per particle, defined as:

$$\mathtt{m} = \left| \frac{1}{N} \sum_{i=1}^{N} s_i \right|, \tag{5.46}$$

and its ensemble average, m,

$$m = \langle \mathtt{m} \rangle \,. \tag{5.47}$$

Averages of any function $G(S)$ are performed using the Boltzmann factor $\mathrm{e}^{-\beta\mathcal{H}}$, or

$$\langle G \rangle = \frac{\sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} G(s_1, \ldots, s_N) e^{-\beta\mathcal{H}}}{\sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-\beta\mathcal{H}}}. \tag{5.48}$$

Other important quantities are the magnetic susceptibility $\chi_T = \left.\dfrac{\partial m(T,H)}{\partial H}\right|_{H=0}$, which can be related to fluctuations of the order parameter:

$$\chi_T = \frac{N}{kT} \sigma^2[\mathtt{m}] = \frac{N}{kT} \left[\langle \mathtt{m}^2 \rangle - \langle \mathtt{m} \rangle^2 \right], \tag{5.49}$$

the internal energy per particle, $u = \mathcal{U}/N$,

$$u = \frac{\langle \mathcal{H} \rangle}{N}, \tag{5.50}$$

the specific heat per particle, $c = C/N$. It can be related to the fluctuations of energy,

$$c = \frac{\sigma^2[\mathcal{H}]}{kT^2 N} = \frac{1}{kT^2 N} \left[\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 \right], \tag{5.51}$$

and many others. Relations of this kind between a response function, χ_T, c, and microscopic fluctuations, $\sigma^2[\mathtt{m}]$, $\sigma^2[\mathcal{H}]$, were first obtained by Einstein and go under the general name of fluctuation-dissipation relations.

The magnetization measures the degree of order: if all spins point in the same direction (either $+1$ or -1) then it takes the maximum value $m = 1$. If the spins point randomly in both directions, then the sum $\sum_{i=1}^{N} s_i$ is close to 0 and, after dividing by N, we get that the magnetization goes to zero as the system size N increases. It is clear from its definition, that the magnetization depends on both the temperature T and the magnetic field H, $m(T, H)$. The value at zero magnetic field,

$m(T,0) \equiv m_0(T)$, is called the spontaneous magnetization. Again, the notation is the one used in magnetic systems: a piece of iron in a magnetic field will display some magnetization and, if the magnetic field is turned off, the magnetization will remain different from zero only below the critical temperature. The detailed calculation for the Ising model shows that this behavior is reproduced depending on the type of lattice. For one-dimensional lattices in which spins have only two nearest neighbors, the case considered by Ising himself, the spontaneous magnetization is always zero, independently of temperature. However, the generic behavior for lattices in two or more dimensions, is the existence of a critical temperature, below which the spontaneous magnetization is indeed different from zero. From the analytical point of view, only a limited set of lattices can be studied, including a variety of two-dimensional lattices and the fully-connected lattice in which every spin is connected to every other spin. It is worth mentioning here the work by Onsager who in an authentic mathematical tour de force, was able to compute the free energy and the related thermodynamic potential in the case of zero magnetic field for the regular square lattice. He was also able to find the spontaneous magnetization. Despite the tremendous difficulty of the calculation, the spontaneous magnetization is given by a deceptively simple expression:

$$m_0(T) = \begin{cases} 0, & T > T_c \\ \left(1 - [\sinh(2J/kT)]^{-4}\right)^{1/8}, & T \leq T_c. \end{cases} \tag{5.52}$$

with a value of the critical temperature $kT_c/J = 2/\log(1+\sqrt{2}) \approx 2.2691853\ldots$.

The alternative to the complicated analytical calculations is the use of the Monte Carlo method. There, we replace the true average by the sample average

$$m = \frac{1}{M}\sum_{k=1}^{M} \mathtt{m}_k, \tag{5.53}$$

being $\mathtt{m}_k$ the value of the magnetization computed in $k = 1,\ldots,M$ spin configurations $S^1,\ldots,S^M$. Remember that each configuration S is a set of values for the N spin variables $S = (s_1,\ldots,s_N)$. The configurations S^k must be generated according to the probability

$$f(S) = \mathcal{Z}^{-1} e^{-\beta\mathcal{H}(S)}. \tag{5.54}$$

Metropolis algorithm

Let us consider first the Metropolis algorithm to generate the configurations S^k. The basic ingredient is the proposal probability $g(S'|S)$. We should not be surprised now of the general strategy: the proposed configuration S' differs from S only in the value of a single spin variable. Therefore, we select a site, say i, and propose a change $s_i \to s_i'$. It should be clear that the only possible proposal is $s_i' = -s_i$, i.e. to propose $s_i' = +1$ if $s_i = -1$ and propose $s_i' = -1$ if $s_i = +1$. Which is the change in the Hamiltonian $\Delta\mathcal{H} = \mathcal{H}(S') - \mathcal{H}(S)$ associated with this proposal? As $N-1$ spins remain unchanged, the only variation in the Hamiltonian comes from

the terms of the sum in (5.45) in which the selected spin s_i appears. Let us denote by $s_{i_1}, \ldots, s_{i_D}$ the set of neighbors of s_i. Therefore, the change is:

$$\Delta\mathcal{H} = \left(-J \sum_{\mu=1}^{D} s_i' s_{i_\mu} - H s_i' \right) - \left(-J \sum_{\mu=1}^{D} s_i s_{i_\mu} - H s_i \right), \tag{5.55}$$

or, using $s_i' = -s_i$,

$$\Delta\mathcal{H} = \left(2J \sum_{\mu=1}^{D} s_{i_\mu} + 2H \right) s_i. \tag{5.56}$$

The acceptance probability can be chosen as $h(S'|S) = \min\left[1, e^{-\beta\Delta\mathcal{H}}\right]$, the Metropolis choice, or $h(S'|S) = \left(1 + e^{-\beta\Delta\mathcal{H}}\right)^{-1}$, Glauber choice, or other convenient expression.

Let us now give some details of the programming in the case of a regular 2-d square lattice. Most of what we will say can be straightforwardly extended to other lattices, regular or random. The first thing we need to do is to store the configuration $(s_1, \ldots, s_N)$. Since the variables are ± 1, it is possible to use very sophisticated storage methods where each variable occupies only one bit of memory (a bit equal to 0 corresponds to $s_i = -1$ and a bit equal to 1 to $s_i = +1$). However, we start by a simple storage method in which spin s_i is stored in location $s(i)$ of an array of integers. We then define `integer s(N)` as the array where we store the variables. Since we are in the square lattice, the number of sites is $N = L^2$, and we could as well store the variables in a two indexes array, such as `integer s(L,L)`. Both notations are equivalent, see figure 5.2. However, the use of a single index brings many simplifications to the program structure and we will keep this way of storing the variables. It is convenient to keep in mind that if we use cartesian coordinates (ix, iy), $ix, iy = 1, \ldots, L$ the single index $i = 1, \ldots, N$ is given by $i = (iy - 1) \times L + ix$, as the reader can check.

In the square lattice, spin s_i has $D = 4$ neighbors than we name $s_{i_1}, s_{i_2}, s_{i_3}, s_{i_4}$. These neighbors are located, respectively, at the right, top, left and bottom sites of i, see figure 5.1. It is convenient, instead of computing every time the indexes i_1, i_2, i_3, i_4 that correspond to site i, to store those values on arrays, such that they can be easily (and quickly) accessed when needed. We use the array `integer n1(N)` such that the value `n1(i)` is equal to i_1, the right neighbor of site i. Similarly, the arrays `integer n2(N), n3(N), n4(N)` store the locations of the up, i_2, left, i_3, and down, i_4, neighbors of site i. A special interest deserve the spins at the borders of the square lattice. One might simply consider that the spins at these borders do not have as many neighbors as the others. A more commonly used choice is that of periodic boundary conditions or pbc, for short. This means that the right neighbor of a site which is located at the right edge of the square is located in the same row at the left edge. And similarly for other directions. Figure 5.3 sketches the nearest-neighbor connectivity in the square lattice. It is as if the right and left edges, and the top and the bottom ones, are connected. The reader might try to imagine how this would look like. The resulting closed structure has the topology of a torus.

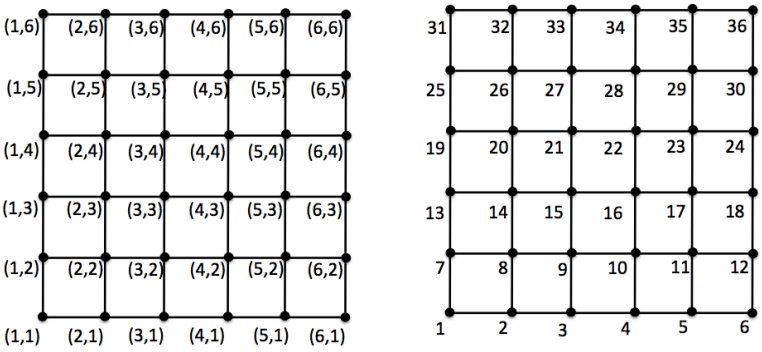

Figure 5.2 The nodes of a square lattice can be stored using two indexes (i_x, i_y), left panel, or a single one i, right panel. The figure exemplifies these two possibilities in the case of a linear side $L = 6$.

In order to generate the right values for those arrays, we use the `subroutine neighbors`. It uses conveniently the equivalence between the one-index i and the two-indexes (i_x, i_y) notation to find the four neighbors of a site. This routine implements the pbc. For example, the right neighbor of site $(i_x = L, i_y)$ is not $(L + 1, i_y)$ but $(1, i_y)$, etc. We urge the reader to have a look at this routine now to understand how it works:

```
subroutine neighbors(n1,n2,n3,n4,L)
dimension n1(L*L),n2(L*L),n3(L*L),n4(L*L)
  do ix=1,L
  do iy=1,L
    i=(iy-1)*L+ix
    ix1=ix+1
      if (ix1.eq.L+1) ix1=1
      n1(i)=(iy-1)*L+ix1
    iy2=iy+1
      if (iy2.eq.L+1) iy2=1
      n2(i)=(iy2-1)*L+ix
    ix3=ix-1
      if (ix3.eq.0) ix3=L
      n3(i)=(iy-1)*L+ix3
    iy4=iy-1
      if (iy4.eq.0) iy4=L
      n4(i)=(iy4-1)*L+ix
  enddo
  enddo
end subroutine neighbors
```

Next thing to consider is the calculation of the acceptance factor. We note that the possible change of energy, as given by (5.56) in the case a magnetic field $H = 0$ can

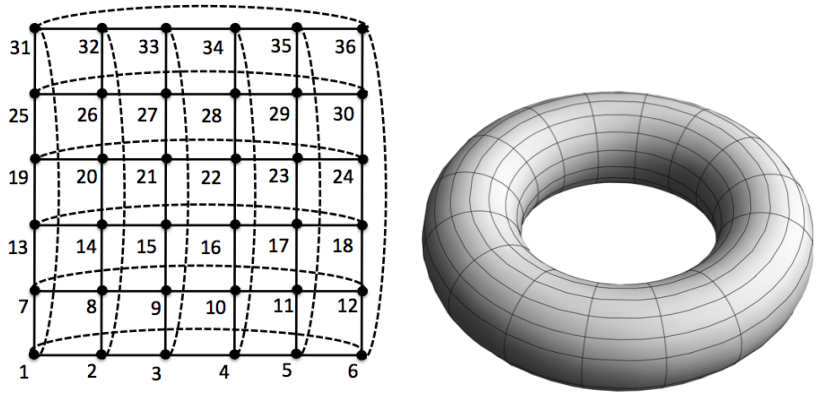

Figure 5.3 The nearest-neighbor connectivity in a square lattice using periodic boundary conditions and the resulting torus topology. For example, the nearest neighbors of site $i = 12$ are $i_1 = 7$, $i_2 = 18$, $i_3 = 11$, $i_4 = 6$, while the nearest neighbors of site $i = 31$ are $i_1 = 32$, $i_2 = 1$, $i_3 = 36$, $i_4 = 25$.

take only 5 possible values, namely:

$$-\beta\Delta\mathcal{H} = -2\beta J s_i \sum_{\mu=1}^{4} s_{i_\mu} = -2KB_i = -2K \begin{cases} -4 \\ -2 \\ 0 \\ 2 \\ 4 \end{cases} \tag{5.57}$$

where we have defined $K = \beta J$ and $B_i = s_i \sum_{\mu=1}^{4} s_{i_\mu}$, being a sum of four numbers, each one being ± 1, can only take the values $-4, -2, 0, 2, 4$, as indicated. Similarly, the Boltzmann-Gibbs factor

$$e^{-\beta\Delta\mathcal{H}} = e^{-2KB_i} \tag{5.58}$$

can only take 5 possible values. We decide to store the acceptance probabilities in the array `h(-4:4)`. This is defined as

$$\mathtt{h(j)} = \min[1, e^{-2Kj}], \qquad j = -4, -2, 0, 2, 4 \tag{5.59}$$

To accept, we compare `h(j)` with a random number u in the usual way. If $u <$`h(j)` we accept the proposal. Otherwise it is discarded and we must make a new proposal by selecting randomly another of the spin variables.

All the preliminary steps are now defined. The program, after creating the arrays of neighbors and setting the initial condition randomly assigning values $s_i = +1$ and $s_i = -1$ with probability $1/2$, then performs the proposal/acceptance steps. These are divided in two blocks: first we perform `M0` $*$ `N` thermalization steps, which is equivalent to `M0` Monte Carlo steps (remember that one MCS is equal to N basic proposal/acceptance steps); next, we begin the `M` measurements. Before each

measurement we perform `mc*N` updates, or `mc`MCS. In this simple implementation we measure three things: the magnetization $\langle \mathtt{m} \rangle$ (stored in `rm`), its fluctuations $\langle \mathtt{m}^2 \rangle - \langle \mathtt{m} \rangle^2$ (stored in `rm2`), and the correlation function $\rho_m(t = \mathtt{mc}\ \text{MCS})$, or the correlation between two consecutive values of the magnetization, (stored in `c`). Finally, the program runs for different values of temperature. In a setup like this, it is convenient to start first at at high value of the temperature ($T = 4$ in the listing below), where the thermalization times are smaller. When lowering the temperature, we do not generate again the initial condition, but use instead the final configuration at the previous, higher, temperature. We now provide a full listing and ask the reader to go through all steps carefully.

```
program Ising_2D_Metropolis
  parameter (L=80,N=L*L)
  implicit double precision(a-h,o-z)
  integer s(N),n1(N),n2(N),n3(N),n4(N)
  dimension h(-4:4)
  data M,M0,mc /8192,1000,1/
  call neighbors(n1,n2,n3,n4,L)
  do i=1,N                              !Initial condition
    if (ran_u().lt.0.5d0) then
      s(i)=+1
    else
      s(i)=-1
    endif
  enddo
  do 999 T=4.0,0.1,-0.1          ! Loop over temperatures
  do j=-4,4,2                    ! Create the array with the
      h(j)=min(1.0,exp(-2*j/T))  ! Boltzmann-Gibbs factors
  enddo
  do  ij=1,M0*N                  ! Thermalizing steps
      i=i_ran(N)
      ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
      if (ran_u().lt.h(ib)) s(i)=-s(i)
  enddo
  c=0.0                         ! Initialize averages
  rm=0.0
  rm2=0.0
  rm1=real(abs(sum(s)))/N
  do  im=1,M                    ! Updating steps
    do ij=1,mc*N
      i=i_ran(N)
      ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
      if (ran_u().lt.h(ib)) s(i)=-s(i)
    enddo
    rm0=real(abs(sum(s)))/N      !  Begin measures
    write(88,*) rm0
    rm=rm+rm0
    rm2=rm2+rm0*rm0
    c=c+rm0*rm1
    rm1=rm0
  enddo
!Final averages
  rm=rm/M
  rm2=rm2/M-rm*rm
```

```
  c=(c/M-rm*rm)/rm2
  if (c.ne.1.0) tau=1.0/(1.0-c)
  error=sqrt(rm2*(2*tau+1)/M)
  write(66,'(f10.6,1p5e16.6)') T,rm,rm2,error,mc*tau,c
999 continue
end program  Ising_2D_Metropolis
```

The kernel of the program are the three lines

```
  i=i_ran(N)
  ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
  if (ran_u().lt.h(ib)) s(i)=-s(i)
```

The first line selects randomly the spin to be updated. The second line computes the necessary index of the acceptance probability. The third line, does the actual updating. In fact, if `h(ib)` ≥ 1, something that occurs if `ib` ≤ 0, there is no need to compare the acceptance probability with a random number, as the proposal is always accepted. This can be implemented by modifying the last line to:

```
  if (ib.le.0) then
   s(i)=-s(i)
  else if (ran_u().lt.h(ib)) s(i)=-s(i)
  endif
```

which reduces the number of calls to the random number generator routine, but increases the complexity of the program. If we would like to implement other updating probabilities, such as Glauber, all we would need to do is to change

```
  do j=-4,4,2
    h(j)=1.0d/(1.0+exp(-2*j/T))
  enddo
```

Another widely used modification is to run sequentially through the sites to update instead of selecting them sequentially. This can be done by choosing the spins in the order $s_1, s_2, s_3, \ldots, s_N$. This requires only a simple modification of the program. Namely, replace the lines around `i=i_ran(N)` by:

```
  do ij=1,mc
   do  i=1,N
      ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
      if (ran_u().lt.h(ib)) s(i)=-s(i)
   enddo
  enddo
```

Another procedure is to divide the lattice in two sub-lattices such that the sites on each lattice do not have neighbors in the same lattice. If the side L is an odd number, this is easily implemented as:

```
  do ij=1,mc
   do  i=1,N,2
      ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
      if (ran_u().lt.h(ib)) s(i)=-s(i)
   enddo
   do  i=2,N,2
```

```
      ib=s(i)*(s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i)))
      if (ran_u().lt.h(ib)) s(i)=-s(i)
   enddo
enddo
```

Finally, note that the program stores in `unit 88` each one of the values of the magnetization. This is necessary, for example, to compute accurately the correlation time τ_m of this quantity using another program, like the one explained in Appendix 14. The program given here uses instead, as a rough estimate for the correlation time, the value $\tau_m = \frac{1}{1-\rho_m(1)}$ derived in (**??**).

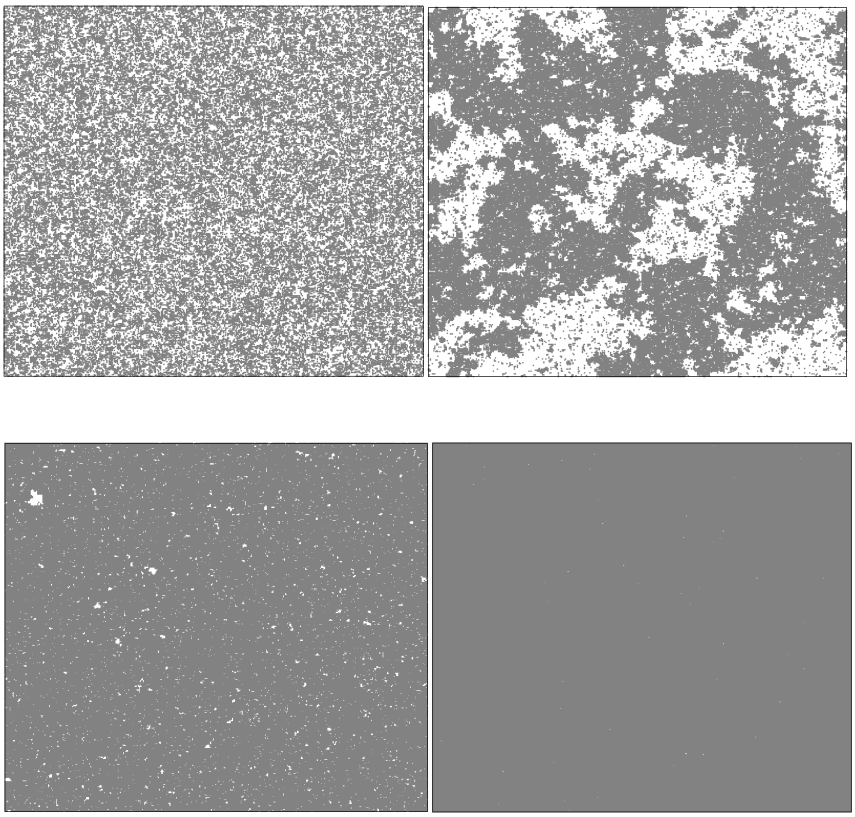

Figure 5.4 Representative configurations of the Ising model at $T = 4$ (top-left), $T = T_c$ (top-right), $T = 2$ (bottom-left), $T = 1$ (bottom-right), for a system of $N = 400^2$ sites. A spin up, $s_i = +1$, is indicated by a dark dot, while a spin down, $s_i = -1$, is represented by a white dot. Observe that at $T = T_c$ the structure is fractal, meaning that there are many blocks of up spins in a background of down spins, and vice versa, recursively.

In figure 5.4 we plot some representative configurations obtained after running this basic Ising model program. We can see the order-disorder phenomenology:

at temperatures larger than the critical temperature T_c, the system is disordered and there is, on a local average, approximately the same number of up and down spins. At low temperatures, one of the two spin options dominates. At $T \approx T_c$, there is still the same average number of up and down spins, but they begin to organize themselves in a fractal-like, self-similar structure, the precursor of the phase transition.

The Kawasaki interpretation of the Ising model

We mentioned already that the Ising model can be used in very different contexts. An interesting interpretation as a model for binary alloy separation was introduced by Kawakasi. In his model, a variable taking a value $s_i = +1$ indicates that on node i there is one atom, A, of a particular metal (e.g. aluminum), while $s_i = -1$ indicates the presence of another type of atom B (e.g. zinc). The Hamiltonian is still given by (5.45) but now the magnetic field H is interpreted as the difference in chemical potential between the two types of atoms. The novelty with respect to the usual interpretation as a ferromagnet, is that atoms A and B can not transmute. This means that the number of A atoms and the number of B atoms are both constant. The averages have to be performed with respect to the same pdf than before, $f = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}}$, but only when the total number of A atoms is a prefixed number N_A (and consequently the number of B atoms is $N_B = N - N_A$.

How do we implement such a constrain in our proposal probability $g(S'|S)$? Kawasaki's proposal is to allow for changes $S \to S'$ in which two neighbor atoms exchange positions. This certainly keeps constant both the number of A and B atoms, they simply move around. This also wants to mimic the real diffusion process that occurs in an alloy as it is cooled down from a high temperature phase. So, in the proposal step, select a pair of neighbor sites i, j, check than $s_i = -s_j$ (otherwise, it does not matter whether we accept the change or not) and propose the change $(s_i, s_j) \to (s'_i, s'_j)$ with $s'_i = s_j = -s_i, s'_j = s_i = -s_j$. It is clear that this is a symmetric proposal $g(S'|S) = g(S|S')$. In order to choose the pair of neighbors sites, we first choose i randomly between 1 and N. Then, in the square lattice, it suffices to select j with equal probability as the neighbor to the right or the neighbor up. Imagine we have selected j as the neighbor to the right, i.e. $s_j = s_{i_1}$. Then, the only terms of the Hamiltonian that will change after exchanging s_i and s_j will be $s_i(s_{i_2} + s_{i_3} + s_{i_4}) + s_j(s_{j_1} + s_{j_2} + s_{j_4})$ (it will help to understand this if the reader plots the lattice and the involved spins). The change in energy will be:

$$\begin{aligned} \Delta\mathcal{H} &= -J\left(s'_i(s_{i_2} + s_{i_3} + s_{i_4}) + s'_j(s_{j_1} + s_{j_2} + s_{j_4})\right) \\ &\quad +J\left(s_i(s_{i_2} + s_{i_3} + s_{i_4}) + s_j(s_{j_1} + s_{j_2} + s_{j_4})\right), \end{aligned} \tag{5.60}$$

or, using that $s'_j = -s'_i = -s_j = s_i$,

$$\Delta\mathcal{H} = 2Js_i\left(s_{i_2} + s_{i_3} + s_{i_4} - s_{j_1} - s_{j_2} - s_{j_4}\right). \tag{5.61}$$

If the site j would have been the one located up of site i, then the charge energy would be similar but involving a different set of sites, namely:

$$\Delta\mathcal{H} = 2Js_i\left(s_{i_1} + s_{i_3} + s_{i_4} - s_{j_1} - s_{j_2} - s_{j_3}\right). \tag{5.62}$$

In any of the two events, one can write $-\beta\Delta\mathcal{H} = -2KB_i$, being B_i a number extracted from the set $(-6, -4, -2, 0, 2, 4, 6)$. The only changes to the above program is to change the dimension of array `h` from `integer h(-4:4)` to `integer h(-6:6)` and not forget to change `dimension h(-6:6)` and write

```
do j=-6,6,2
  h(j)=min(1.0,exp(-2*j/T))
enddo
```

The updating part is

```
1 i=i_ran(N)
  if (ran_u().lt.0.5d0) then
   j=n1(i)
   if (s(i).eq.s(j)) goto 1
   ib=s(i)*(s(n2(i))+s(n3(i))+s(n4(i))-s(n1(j))-s(n2(j))-s(n4(j)))
  else
   j=n2(i)
   if (s(i).eq.s(j)) goto 1
   ib=s(i)*(s(n1(i))+s(n3(i))+s(n4(i))-s(n1(j))-s(n2(j))-s(n3(j)))
  endif
  if (ran_u().lt.h(ib)) then
     s(i)=-s(i)
     s(j)=-s(j)
   endif
```

Of course, now it does not make sense to compute the average of the magnetization, as this is a constant (it is the number of A atoms minus the number of B atoms). A typical quantity to measure instead is the number of links that join two nearest neighbor sites in which the variables take different values, a sort of contact area between the metals of the two alloys.

One of the most important drawbacks of this algorithm is that it might occur too often that the two selected sites hold the same type of atoms, and then this trial has to be repeated. A possibility is to make a list of all possible pairs of neighboring sites that hold different values for the spin variables and then choose a pair only from this list. This possibility speeds up the program but it does require some amount of programming[8)]

Considered as a real model for a binary alloy, the Kawasaki model is able to reproduce some well known facts known to metallurgists. First, the process leads to phase separation (with each alloy occupying mainly a localized spatial area) only below a critical temperature, T_c. For temperatures larger than T_c the two metals are well mixed, while for a temperature smaller than T_c, the system splits in two phases, each one of them rich in one of the metals. The second interesting feature of Kawasaki model is that the time evolution of the phase separation process occurring at temperatures below T_c depends strongly on the initial relative proportion of each metal. If this proportion is close to 50%, then the process proceeds by what is called "spinodal decomposition" in which filament-like structures linking all the atoms of

8) This modification constitutes the *n-fold way* proposed by Bortz, Kalos and Lebowitz. It will be met again in a different context when we discuss the numerical integration of master equations in chapter 9.

one metal begin to form and then coarsen and grow. When one of the metals is more abundant than the other, then the metal which is in the minority starts by forming small droplets embedded in the majority metal. These droplets then begin to coalescence and form larger droplets, until only one big one (and maybe several very minor droplets) can be seen. This evolution process can be observed in figure 5.5.

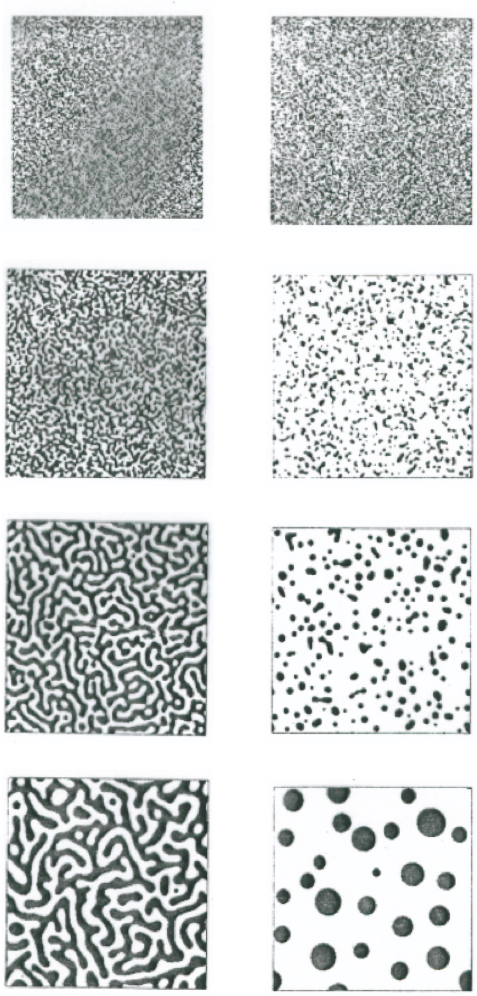

Figure 5.5 Representative configurations of the Kawasaki version of the Ising model at $T = 0.95$ starting from an initial condition with exactly half of the spins in the $+1$ state (left panel), and starting with 1/3 of the spins in the $+1$ state (right panels). In both cases, time runs from top to bottom panels. In the left panels we see the spinodal decomposition evolution mechanism, while in the right we observe the nucleation and growth of droplets of one phase.

Heat bath algorithm

The Ising model (we now return to the magnetic version in which the number of spins of one or the other type does not need to be conserved) offers an interesting and relatively simple application of the heat-bath method. Remember that in this method, the acceptance probability is always equal to one, but only after the proposal probability has been carefully chosen (it is not arbitrary anymore) in order to ensure that the detailed balance condition is satisfied. So, we want to propose the change from configuration $S = (s_1, \ldots, s_i, \ldots, s_N)$ to $S' = (s_1, \ldots, s'_i, \ldots, s_N)$ in which the only change is in one spin located at, say, site i. According to (4.77) the proposal probability $g(S'|S) \equiv g(s'_i)$ is given by the conditional probability:

$$\begin{aligned} g(s'_i) &= f(s'_i | s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N) \\ &= \frac{f(s_1, \ldots, s_{i-1}, s'_i, s_{i+1}, \ldots, s_N)}{f(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)} \end{aligned} \tag{5.63}$$

This might look as a very complicated expression (and maybe it is so), but we must take into account that we are proposing a new value only for s'_i, while all other variables remain unchanged. All we have to do, then, is to look for the dependence on s'_i of the previous expression, all other variables take simply fixed numbers, in other words, they are constants. If we look at the numerator of the last formula and use $f = e^{-\beta \mathcal{H}}$ we realize that the only place where s'_i appears is in the term in $\mathcal{H}$ in which it is multiplied to its neighbors, namely $\sum_\mu s'_i s_{i_\mu}$, being s_{i_μ} the set of D neighbors of s_i ($D = 4$ neighbors in the square lattice), therefore, we have:

$$g(s'_i) = C^{-1} e^{\beta J s'_i \sum_\mu s_{i_\mu}} \equiv C^{-1} e^{a_i s'_i} \tag{5.64}$$

where we have defined $a_i = \beta J \sum_\mu s_{i_\mu}$ and C is the normalization constant, hiding all the nasty dependence in the rest of the spin variables. The nice part is that, whatever complicated the expression for C is if we write it in full, we have a much simpler method to compute its actual value. Simply realize that s'_i can only take the values $s'_i = \pm 1$ and hence the probabilities of these two values must add up to one, $g(s'_i = +1) + g(s'_i = -1) = 1$ which leads immediately to $C = e^{a_i} + e^{-a_i}$, or:

$$g(s'_i = +1) = \frac{e^{a_i}}{e^{a_i} + e^{-a_i}} = \frac{1}{1 + e^{-2a_i}}, \tag{5.65}$$

$$g(s'_i = -1) = \frac{e^{-a_i}}{e^{a_i} + e^{-a_i}} = \frac{e^{-2a_i}}{1 + e^{-2a_i}}. \tag{5.66}$$

The rest is simple, generate a random number u uniformly distributed in $(0, 1)$. If $u < g(s'_i = +1)$ set $s'_i = 1$, otherwise set $s'_i = -1$. This new value is always accepted. Again we can use the fact that the sum $\sum_\mu s_{i_\mu}$ can only take values from a limited set, namely $(-4, -2, 0, 2, 4)$ in the square lattice, which allows us to compute and store the possible values of $g(s'_i)$. Consequently, we define `dimension g(-4:4)` and fill up the elements of this array with the corresponding values of $g(s'_i = +1)$, namely

```
do j=-4,4,2
  g(j)=1./(1.+exp(-2.*j/T)
enddo
```

and then the updating steps are

```
do ij=1,mc*N
  i=i_ran(N)
  ia=s(n1(i))+s(n2(i))+s(n3(i))+s(n4(i))
  if (ran_u().lt.g(ia)) then
    s(i)=+1
  else
    s(i)=-1
  endif
enddo
```

5.5 Heisenberg model

This is also a lattice model, but the variable associated to lattice site i is a D-dimensional vector $\vec{s}_i$ of modulus one $|\vec{s}_i| = 1$. The Hamiltonian is similar to that of the Ising model but it involves a scalar product:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{s}_i \cdot \vec{s}_j - \vec{H} \cdot \sum_i \vec{s}_i, \tag{5.67}$$

being $\vec{H}$ the external magnetic field. It is important not to mislead the spatial dimension d of the lattice with the dimension D of the spins on the lattice. We now explain a heat-bath method in order to generate configurations $X = (\vec{s}_1, \ldots, \vec{s}_N)$ distributed according to the Boltzmann weight $f(S) = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}}$. We restrict ourselves to the D=3 case, but appropriate extensions are possible to other dimensions. We do not go through all the details which are very similar to the Ising model case. What we have to do now is to propose a value of a vector $\vec{s}_i\,'$ chosen from the distribution

$$g(\vec{s}_i\,') = C^{-1} e^{\beta (J \sum_\mu \vec{s}_{i_\mu} + \vec{H}) \cdot \vec{s}_i\,'} \equiv C^{-1} e^{\vec{a}_i \cdot \vec{s}_i\,'} \tag{5.68}$$

where C is a normalization constant and we have defined $\vec{a}_i = \beta(J \sum_\mu \vec{s}_{i_\mu} + \vec{H})$. In the heat-bath method, we need to generate a value of $\vec{s}_i\,'$ extracted from the distribution $g(\vec{s}_i\,')$ assuming that the "local field" $\vec{a}_i$ (which contains information about the other spins) is a constant.

Now, a three-dimensional vector of modulus 1 can be defined by giving the two orientation angles $\vec{s}_i\,' = (\phi_i, \theta_i)$ with $\phi_i \in (0, 2\pi)$ and $\theta_i \in (0, \pi)$. The two angles (ϕ_i, θ_i) can be taken with respect to any orientation and it makes sense, according to the expression above, to consider them with respect to the orientation given by the local field $\vec{a}_i$. Furthermore, it is convenient to work with a new variable defined as $\xi_i = \cos(\theta_i)$ which takes values in the interval $\xi_i \in (-1, 1)$, such that the scalar product becomes $\vec{a}_i \cdot \vec{s}_i\,' = a_i \xi_i$. It is possible to find the constant C from the normalization condition:

$$\int_{-1}^{+1} d\xi_i \int_0^{2\pi} d\phi_i g(\phi_i, \xi_i) = \int_{-1}^{+1} d\xi_i \int_0^{2\pi} d\phi_i C^{-1} e^{a_i \xi_i} = 1, \tag{5.69}$$

or

$$C = 2\pi \frac{e^{a_i} - e^{-a_i}}{a_i}, \tag{5.70}$$

yielding a proposal pdf wich can be split as $g(\phi_i, \xi_i) = g(\phi_i) \times g(\xi_i)$, with $g(\phi_i) = 1/2\pi$ a uniform distribution in the interval $(0, 2\pi)$ and

$$g(\xi_i) = \frac{a_i}{e^{a_i} - e^{-a_i}} e^{a_i \xi_i}, \quad \xi_i \in (-1, 1). \tag{5.71}$$

The angle ϕ_i is simply generated by $\phi_i = 2\pi u_i$, being u_i a $\hat{\mathbf{U}}(0, 1)$, a random number distributed uniformly in the interval$(0, 1)$. Concerning $g(\xi_i)$ it can be implemented using the general inversion method, or solving

$$v_i = \int_{-1}^{\xi_i} d\xi_i' g(\xi_i') \tag{5.72}$$

being v_i an independent $\hat{\mathbf{U}}(0, 1)$ number. Solving this equation we obtain:

$$\xi_i = \frac{1}{a_i} \ln \left(1 + (e^{2a_i} - 1)v_i\right) - 1. \tag{5.73}$$

Once ϕ_i and ξ_i have been obtained, it is a matter of algebra to find the expression of $\vec{s_i}\,'$ relative to whatever fixed coordinate system. Typically, the external field $\vec{H}$ serves to define the Z direction.

5.6 Lattice Φ^4 Model

As a final example of the application of the Monte Carlo methods in statistical physics, we consider the so-called lattice Φ^4 model. The model is the lattice version of a field model. By a field model we mean that in every point of space $\vec{r}$ there is defined a real variable, $\Phi(\vec{r})$. In the lattice version, there are only variables Φ_i defined in the sites $i = 1, \ldots, N$ of a lattice. At variance with the Ising or Heisenberg models, where some restriction applied to the site variable (either a binary variable or a vector of modulus 1), in this model each variable can take any real value.

In the field version, the Hamiltonian is a functional of the field, namely:

$$\mathcal{H}(\{\Phi(\vec{r})\}) = \int d\vec{r} \left[\frac{-B}{2} \Phi(\vec{r})^2 + \frac{U}{4} \Phi(\vec{r})^4 + \frac{K}{2} |\vec{\nabla} \Phi(\vec{r})|^2 - H(\vec{r}) \Phi(\vec{r}) \right], \tag{5.74}$$

being B, U and K parameters of the model and $H(\vec{r})$ an external field, usually assumed to be constant $H(\vec{r}) = H$. This field model has been used in many different contexts, in such different topics as quantum field theory or in the study of phase transitions of interest here. One can think, for instance, on a fluid and then the field $\Phi(\vec{r})$ is the density field. Of course, due to the atomic nature of matter, it is difficult to define a density field at every point of space as this would be a highly discontinuous function. One can then parcel the space in cells forming a lattice and

define in each lattice site i the variable Φ_i as the average density in the cell associated to the site. The continuous, field model, interpretation is preferred in the analytical calculations despite some problems that appear at very small spatial scales (reflecting precisely the discontinuity of matter at the atomic scale). It has been extensively used in renormalization group studies in momentum space. The lattice version is the preferred one for numerical calculations and provides as well a regularization of the Hamiltonian of the continuous model.

For the lattice version, we consider a d-dimensional (hyper-cubic) regular lattice Λ, consisting of $N = L^d$ points. Every point $i = 1, \ldots, N$ of this lattice has d coordinates: $i = (j_1, \ldots, j_d)$. The set of $2d$ neighbors have coordinates

$$\begin{aligned}
&i_1 = (j_1 + 1, \ldots, j_d),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&i_d = (j_1, \ldots, j_d + 1),\\
&i_{d+1} = (j_1 - 1, \ldots, j_d),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&i_{2d} = (j_1, \ldots, j_d - 1).
\end{aligned}$$

Periodic boundary conditions are assumed on this lattice. On every site of the lattice there is a scalar variable Φ_i. The set of all variables is $[\Phi] \equiv (\Phi_1, \ldots, \Phi_N)$. We also introduce a Hamiltonian function $\mathcal{H}$ given by:

$$\mathcal{H}([\Phi]) = \sum_{i=1}^{N} a_0^d \left[\frac{-B}{2}\Phi_i^2 + \frac{U}{4}\Phi_i^4 + \frac{K}{2}\sum_{\mu=1}^{d}\left(\frac{\Phi_{i_\mu} - \Phi_i}{a_0}\right)^2 - H\Phi_i \right], \tag{5.75}$$

where the parameter a_0 (the lattice spacing) has been introduced to stress the fact that the continuous, field, version can be recovered (if needed) taking the limit $a_0 \to 0$ and $L \to \infty$. Note the replacement in (5.75) of the integral of (5.74) by a sum and the spatial derivative of the gradient by a difference. If we are not interested in the continuous version we can simply set $a_0 = 1$ as it only implies a rescaling of parameters B, U, K and H. In the applications to the studies of phase transitions, the parameter B depends on temperature $B(T)$. The first two and the last terms of the sum appearing in the Hamiltonian are local terms (depending only on the field at location i) and can be thought of as local potential terms $V(\Phi_i)$:

$$V(\Phi_i) = \frac{-B}{2}\Phi_i^2 + \frac{U}{4}\Phi_i^4 - H\Phi_i. \tag{5.76}$$

The third term in the Hamiltonian (5.75), the one multiplied by K, is called the interaction term as it contains cross terms between fields in nearest-neighbor sites in the regular lattice. Expanding the square and noticing that:

$$\sum_{i=1}^{N}\sum_{\mu=1}^{d}\Phi_{i_\mu}^2 = d\sum_{i=1}^{N}\Phi_i^2, \tag{5.77}$$

it is possible to rewrite the Hamiltonian (after setting $a_0 = 1$) as:

$$\mathcal{H}([\Phi]) = \sum_{i=1}^{N}\left[\frac{-B + 2dK}{2}\Phi_i^2 + \frac{U}{4}\Phi_i^4 - K\sum_{\mu=1}^{d}\Phi_{i_\mu}\Phi_i - H\Phi_i \right]. \tag{5.78}$$

Note that the interaction terms contain all the products of nearest-neighbor variables, leading to a form which is not so different from that of the Ising model:

$$\mathcal{H}([\Phi]) = \sum_{i=1}^{N} \left[\frac{-B + 2dK}{2} \Phi_i^2 + \frac{U}{4} \Phi_i^4 - H\Phi_i \right] - K \sum_{\langle i,j \rangle} \Phi_i \Phi_j \,. \tag{5.79}$$

From a statistical mechanics analysis of this model, the importance of the Hamiltonian lies in that it dictates the pdf of the set of variables $\mathbf{\Phi} \equiv (\Phi_1, \ldots, \Phi_N)$ as

$$f(\mathbf{\Phi}) = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}(\mathbf{\Phi})}, \tag{5.80}$$

being $\mathcal{Z}$ the partition function, computed as the multiple integral of the Boltzmann factor for all the field values:

$$\mathcal{Z} = \int_{-\infty}^{\infty} d\Phi_1 \ldots \int_{-\infty}^{\infty} d\Phi_N e^{-\beta \mathcal{H}([\Phi])} \equiv \int d\mathbf{\Phi} e^{-\beta \mathcal{H}([\Phi])}. \tag{5.81}$$

For this integral to exist, i.e. for the Φ^4 model to be consistent, it is required that the parameter U verifies $U > 0$.

As usual, the magnitudes of interest are computed as averages of field functions, $G(\mathbf{\Phi})$, with the probability density function $f(\mathbf{\Phi})$:

$$\langle G(\mathbf{\Phi}) \rangle = \mathcal{Z}^{-1} \int d\mathbf{\Phi} G(\mathbf{\Phi}) e^{-\beta \mathcal{H}(\mathbf{\Phi})}. \tag{5.82}$$

Example of quantities of interest are the *magnetization*, m defined as:

$$m = \left\langle \left| \frac{1}{N} \sum_{i=1}^{N} \Phi_i \right| \right\rangle, \tag{5.83}$$

and, particularly, its value at $H = 0$, the spontaneous magnetization m_0. Another quantities of interest are the internal energy, $\mathcal{U}$:

$$\mathcal{U} = \langle \mathcal{H}(\mathbf{\Phi}) \rangle \,, \tag{5.84}$$

as well as the magnetic susceptibility, χ_T and the specific heat, defined as the fluctuations of the magnetization and the internal energy, using the same definitions (5.49)-(5.51) than in the Ising model.

Before explaining the specificity of the numerical methods of the Monte Carlo type to compute the above averages, we shall briefly review some qualitative aspects as given by a simple approximate theory, known as mean-field approximation and first introduced by Landau. The basic idea is to assume that the different variables Φ_i adopt a common value Φ_0, the mean field. The common value Φ_0 is chosen, logically, as the one that maximizes the pdf (5.80) or, equivalently, as the one that minimizes the Hamiltonian $\mathcal{H}$. After setting $\Phi_i = \Phi_0$, $\forall i$ in (5.75) the term proportional to K vanishes and all the terms in the sum are identical, leading to $\mathcal{H} = NV(\Phi_0)$. Hence, the minima of $\mathcal{H}$ are those of the local potential $V(\Phi_0)$. Let us consider first the case of no external field, $H = 0$. In this case, the local potential changes qualitative when

the parameter B changes sign. If $B < 0$ the local potential has only one minimum at $\Phi_0 = 0$. On the other hand, when $B > 0$ there are two minima of equal depth located at $\Phi_0 = \pm\sqrt{B/U}$. As the magnetization is defined as the average of the absolute value of the field, it follows the prediction of the mean-field theory for the spontaneous magnetization:

$$m_0 = \begin{cases} 0, & B < 0, \\ \sqrt{\frac{B}{U}}, & B \geq 0, \end{cases} \tag{5.85}$$

As said before, the parameter B is assumed to be a decreasing function of temperature, such as $B(T) = B_0(T_0 - T)$, for example. The condition $B(T_0) = 0$ defines the critical temperature T_0. The magnetization is then zero for temperatures above T_0 and non-zero below T_0. This is the basic phenomenology of phase transitions and allow us to interpret `m` as the order parameter.

The basic picture given by the mean-field approximation is qualitatively correct[9], but not the details. It is still true that the magnetization is zero below some critical temperature T_c, but this is not determined by the condition $B(T_c) = 0$ and, instead,we find $B_c = B(T_c) \neq 0$. This can be summarized as:

$$m_0 = \begin{cases} 0, & B < B_c, \\ m_0(B) \neq 0, & B \geq B_c, \end{cases} \tag{5.86}$$

being B_c dependent on the other parameters of the theory, U and K. The function $m(B)$ can be expanded near $B \gtrsim B_c$ as $m(B) = m_0(B - B_c)^\beta$, being m_0 a constant, and β the critical exponent. The determination of critical exponents is one of the main points of interest in the modern theory of phase transitions. It is proven that the different phase transitions can be classified in different "universality classes", each class characterized, amongst other things, by a common value of the critical exponents. In particular, it is known that the Ising and the Φ^4 models at the same spatial dimension belong to the same universality class.

Monte Carlo methods

After this short introduction to the phenomenology of the Φ^4 model, let us now explain the implementation of Monte Carlo methods for the numerical calculation of averages such as the magnetization `m` of the internal energy $\mathcal{U}$. It is convenient to start first with a simplification of parameters. Let us redefine:

$$\phi_i = (\beta K)^{1/2}\Phi_i, \tag{5.87}$$

$$b = \frac{B}{K} - 2d, \tag{5.88}$$

$$u = \frac{U}{\beta K^2}, \tag{5.89}$$

$$h = H\left(\frac{\beta}{K}\right)^{1/2}, \tag{5.90}$$

9) To be precise, it is correct if the spatial dimension d is larger than 1.

a rescaling chosen such that

$$\beta\mathcal{H}(\phi) = \sum_{i=1}^{N}\left[\frac{-b-2d}{2}\phi_i^2 + \frac{u}{4}\phi_i^4 + \frac{1}{2}\sum_{\mu=1}^{d}\left(\phi_{i_\mu} - \phi_i\right)^2 - h\phi_i\right], \tag{5.91}$$

or, expanding the squares as we did before:

$$\beta\mathcal{H}(\phi) = \sum_{i=1}^{N}\left[\frac{-b}{2}\phi_i^2 + \frac{u}{4}\phi_i^4 - h\phi_i\right] - \sum_{\langle i,j\rangle}\phi_i\phi_j. \tag{5.92}$$

The pdf for the new rescaled variables is:

$$f(\phi) = Z^{-1}e^{-\beta\mathcal{H}(\phi)}, \tag{5.93}$$

with Z the normalization constant. In this reparametrization, we are left with only 3 parameters: b, u and h.

In order to generate representative configurations $\phi = (\phi_1, \ldots, \phi_N)$ distributed according to this pdf, we can use a simple[10] Metropolis algorithm. Propose a new configuration in which only one variable, say ϕ_i, is proposed to change $\phi_i \to \phi_i'$ and all other variables remain unchanged. A simple choice is to take ϕ_i' uniformly from the interval $(\phi_i - \Delta, \phi_i + \Delta)$, being Δ a parameter to be chosen to optimize the method. The proposal ϕ_i' is accepted with a probability

$$h(\phi'|\phi) = \min(1, e^{-\beta\mathcal{H}}). \tag{5.94}$$

In computing the change in the Hamiltonian implied by the proposed change, we do not need to use the full expression (5.92), but rather notice that most of the terms disappear when subtracting the old and the new values, such that the change is simply:

$$\beta\Delta\mathcal{H} = \frac{b}{2}({\phi'}_i^2 - \phi_i^2) + \frac{u}{4}({\phi'}_i^4 - \phi_i^4) - ({\phi'}_i - \phi_i)\left(h + \sum_{\mu=1}^{2d}\phi_{i_\mu}\right). \tag{5.95}$$

Alternatively, one could devise a heat-bath method. First, we select a site i and want to propose a value ϕ_i for the variable at that site. This is done with a proposal $g(\phi_i)$ which is proportional to $e^{-\beta\mathcal{H}}$ but where all other variables ϕ_j, $j \neq i$ are considered to be constants. All we need to consider when writing down $g(\phi_i)$ are those terms in the Hamiltonian where ϕ_i appears explicitly. After these considerations, we have:

$$g(\phi_i) = C^{-1}e^{-\frac{u}{4}\phi_i^4 + \frac{b}{2}\phi_i^2 + b_i\phi_i} \tag{5.96}$$

with $b_i = \left(h + \sum_{\mu=1}^{2d}\phi_{i_\mu}\right)$, and C is the normalization constant:

$$C = \int_{-\infty}^{\infty} d\phi_i\, e^{-\frac{u}{4}\phi_i^4 + \frac{b}{2}\phi_i^2 + b_i\phi_i}. \tag{5.97}$$

10) We hope that, by now, the reader finds "simple" this kind of techniques.

We can use now any method to sample the one-variable pdf $g(\phi_i)$. As the constant C might be cumbersome to obtain in some cases, it is convenient to use a rejection method to sample $g(\phi_i)$. For instance, we could split

$$g(\phi) \propto \frac{1}{\sigma\sqrt{2\pi}} e^{-(\phi-\mu)^2/2\sigma^2} \times e^{-\frac{u}{4}\phi^4} \tag{5.98}$$

with $\sigma^2 = -1/b$, a procedure only valid for $b < 0$ (or $B < 2dK$ in the original parameters of the model) and $\mu = b_i/b$. Then, we could use a rejection method, consisting in proposing a value of ϕ_i according to a Gaussian distribution of mean μ and variance σ^2 and then accept that proposed value with a probability $e^{-\frac{u}{4}\phi_i^4} \in (0, 1)$. This could be programmed with the following lines:

```
1   phi=sigma*ran_g()+mu
    if (ran_u().gt.exp(-0.25*u*phi**4) goto 1
```

Remember that in the heat-bath method, the proposals are always accepted.

Many other choices for the proposal and acceptance probabilities are, of course, possible. A characteristic feature of the Φ^4 model, present as well in many other models, is that it is possible to split the Hamiltonian into local and interaction terms. We have already explained that local terms depend only on one of the variables, while interaction terms depend on several variables. This is precisely what we had done in (5.92) that we rewrite as

$$\mathcal{H} = \sum_{i=1}^{N} v(\phi_i) + \mathcal{H}_I \tag{5.99}$$

with

$$\beta v(\phi_i) = -\frac{b}{2}\phi_i^2 + \frac{u}{4}\phi_i^4 - h\phi_i, \tag{5.100}$$

$$\beta \mathcal{H}_I = -\sum_{\langle i,j \rangle} \phi_i \phi_j. \tag{5.101}$$

We now use an approach in which the proposal depends only on the local term, while the acceptance probability will take care of the interaction terms. Specifically, we use a proposal $g(\phi'|\phi)$ which depends only on one variable selected randomly. Once this variable, say ϕ_i, has been chosen, we use a proposal proportional to the Boltzmann factor of the local term, namely

$$g(\phi_i) = C^{-1} e^{-\beta v(\phi_i)}, \tag{5.102}$$

being C a normalization constant. Replacing this proposal in the detailed balance condition (5.17), we obtain

$$e^{-\beta v(\phi_i')} h(\phi'|\phi) e^{-\beta v(\phi_i)} e^{-\beta \mathcal{H}_I(\phi)} = e^{-\beta v(\phi_i)} h(\phi|\phi') e^{-\beta v(\phi_i')} e^{-\beta \mathcal{H}_I(\phi')} \tag{5.103}$$

or

$$h(\phi'|\phi) e^{-\beta \mathcal{H}_I(\phi)} = h(\phi|\phi') e^{-\beta \mathcal{H}_I(\phi')}. \tag{5.104}$$

We have seen these functional equations before. They have the same form that the general detailed balance condition under the symmetric proposal (when the proposal functions g disappear from the equation). The novelty now is that only the interaction part of the Hamiltonian appears in both sides of this equation. Possible solutions are the Metropolis solution:

$$h(\phi'|\phi) = \min(1, e^{-\beta\Delta\mathcal{H}_I}), \tag{5.105}$$

the Glauber solution

$$h(\phi'|\phi) = \frac{e^{-\beta\Delta\mathcal{H}_I}}{1 + e^{-\beta\Delta\mathcal{H}_I}}, \tag{5.106}$$

etc.

Again, when computing $\Delta\mathcal{H}_I$ it is important to realize that this difference only depends on a reduced number of variables. For example, for the Φ^4 model we would propose a value ϕ_i' sampled from the distribution

$$g(\phi_i') = C^{-1} e^{\frac{b}{2}\phi_i^2 - \frac{u}{4}\phi_i^4 + h\phi_i}. \tag{5.107}$$

Instead of using a rejection method as we proposed to sample (5.96), we can sample this new $g(\phi_i')$ using, for example, a numerical inversion algorithm, as, at variance with (5.96), the parameters of (5.107), b, u, h, do not vary after each step. The change in the interaction Hamiltonian needed for the acceptance step is then:

$$\Delta\mathcal{H}_I = (\phi_i' - \phi_i) \sum_{\mu=1}^{d} \phi_{i_\mu} \tag{5.108}$$

In general, a procedure similar to this one can be used when the Hamiltonian can be split in a sum of local terms plus an interaction term. One can choose the new value of the variable, independently of the old value, according to the distribution dictated by the local term. This proposal is then accepted with the Metropolis probability using only the interaction term.

5.7 Data analysis: Problems around the critical region

Once we have developed our Monte Carlo code, the next logical step is to run it and to obtain some results of interest. However, this process is not exempt of danger as we will see using as an example the Ising model in the square lattice. In figure 5.6 we plot the spontaneous (at zero magnetic field) magnetization of the Ising model as obtained from our program for three different system sizes: $L = 20,\ 40,\ 80$, as well as the exact solution given by Onsager, (5.52).

From this figure we notice two clear things: (1) for a considerable range of temperatures, the numerical plots are well outside the theoretical result and (2) the error bars are not constant for all values of the temperature, but they are considerable larger near the critical region. The first discrepancy is due to finite-size effects. Concerning

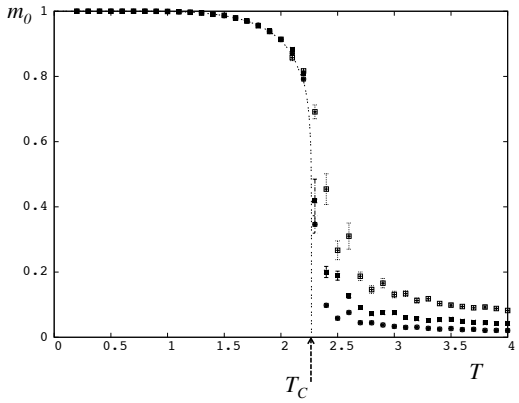

Figure 5.6 Spontaneous magnetization m_0 of the Ising model in the square lattice as a function of temperature T for different values of the system side $L = 20,\ 40,\ 80$ from top to bottom in the right hand side of the figure. The dots have been obtained using the Metropolis algorithm described in the main text. The dashed line is Onsager's exact solution, as given by Eq. (5.52). We have set units such that $k = J = 1$.

the magnitude of the error bars, we recall that the error in the sample average of the magnetization $\mathtt{m}$ is given by (5.14) that we rewrite here:

$$\epsilon[\mathtt{m}] = \frac{\sigma[\mathtt{m}]}{\sqrt{M}}\sqrt{2\tau_{\mathtt{m}} + 1} \tag{5.109}$$

being M the number of points contributing to the sample average, $\sigma^2[\mathtt{m}] = \langle \mathtt{m}^2 \rangle - \langle \mathtt{m} \rangle^2$ the variance of m and $\tau_{\mathtt{m}}$ the correlation time associated to the normalized correlation function $\rho_{\mathtt{m}}$. Certainly the errors decrease by increasing the number of measurements, M, and we can ask which one of the two factors, $\sigma[\mathtt{m}]$ or $\tau_{\mathtt{m}}$, is responsible for the observed error increase around the critical region. The not so encouraging answer is that both factors contribute. Around the critical region one observes: (i) an increase of fluctuations $\sigma[\mathtt{m}]$ and (ii) an increase of the correlation time $\tau_{\mathtt{m}}$ due to the so-called critical slowing down. Finite size effects, increase of fluctuations and critical slowing down and thermalization are the major points of concern in many Monte Carlo simulations.We now discuss these points separately First of all, we must say that these problems are not specific of the Ising model. They can also be observed in the numerical simulations of the Φ^4 model, see figure 5.7 and of every other system that we simulate near a phase transition.

Finite-size effects Finite-size effects appear as the discontinuities and mathematical singularities of a true phase transition requires of the thermodynamic limit $N \to \infty$. For example, Onsager's solution (5.52) implies a discontinuity of the derivative of $m_0(T)$ at $T = T_c$. In fact, the theory of phase transitions predicts that the

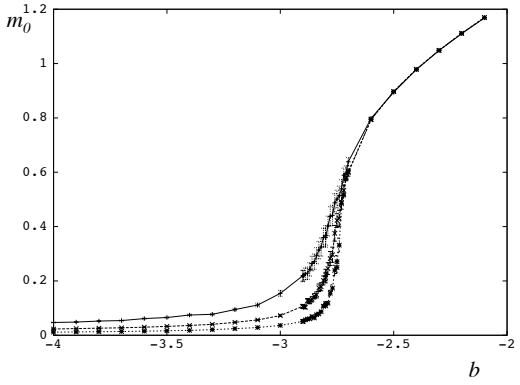

Figure 5.7 Spontaneous magnetization of the Φ^4 model in the square lattice as a function of the parameter b for $u = 1$ and different values of the system side $L = 20, 40, 80$ (from bottom to top lines). The data have been generated using the heat-bath Monte Carlo method explained in the main text.

asymptotic behavior near the critical point is:

$$m_0(T) \sim |1 - T/T_c|^{\beta}, \quad \text{for } T \leq T_c, \tag{5.110}$$

with a value of the "critical exponent" $\beta = 1/8$ for the $2d$ Ising model. Similarly, it predicts that the magnetic susceptibility (5.49) diverges at the critical point as:

$$\chi_T(T) \sim |1 - T/T_c|^{-\gamma} \tag{5.111}$$

with a value of the critical exponent $\gamma = 7/4$, for the 2-d Ising model, implying, formally, that the magnetic susceptibility diverges at $T = T_c$, and, according to (5.49), that the variance of the order parameter greatly increases near T_c. These are examples of the typical non-analytical behavior that occurs at the critical point. Strictly speaking, however, a non-analytical behavior can not appear in a finite system[11]. What it is observed in the simulations, though, is a big increase of the susceptibility in the neighborhood of the critical region, see figure 5.8. As the system size N increases, the maximum in the susceptibility grows higher and its location approaches the true critical temperature T_c of the infinite system.

It can be understood intuitively why finite size effects will be more important near a second order phase transition. In this situation the correlation length, which measures the linear range over which spins at different sites of the lattice are correlated, diverges (in an infinite system) with a power-law singularity:

$$\xi(T) \sim |1 - T/T_c|^{-\nu}, \tag{5.112}$$

11) Think in terms of our simulation. It is impossible that χ_T is equal to ∞ anywhere as the fluctuations of the order parameter are bounded, $\mathtt{m} < 1$, and we never divide by 0.

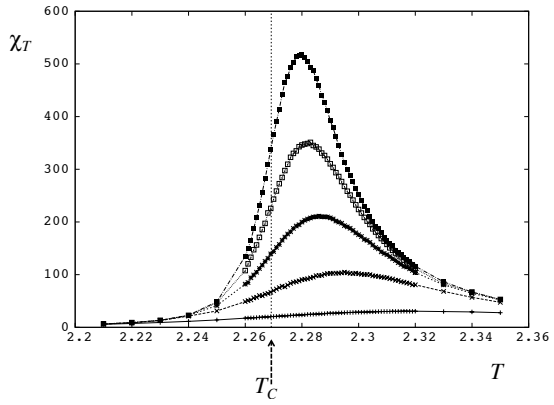

Figure 5.8 Magnetic susceptibility χ_T for the 2-d square Ising model computed from (5.49) using the Metropolis algorithm for system sides $L = 40, 80, 120, 160, 200$. The maximum value increases for increasing L.

being $\nu = 1$ for the 2d Ising model. For a finite system, the correlations can not extend longer than the system side and we must have $\xi \sim L$. The theory of finite size tells us exactly how (and why!) the averages of interest behave. The basic idea is that now the magnetization becomes a *homogeneous* function of ξ and the system side L, $m(\xi, L) = \xi^x \tilde{m}(\xi/L)$. The unknown exponent x is obtained by demanding that in the infinite system, and close enough to the critical point, one recovers the known behavior given by (5.110). This implies that the function $\tilde{m}(z)$ takes a finite limit when $z \to 0$ and then:

$$m_0(T) = \lim_{L\to\infty} m(T, L) = \xi^x \tilde{m}(0) \sim \left[|1 - T/T_c|^{-\nu}\right]^x \sim |1-T/T_c|^{-x\nu}. \quad (5.113)$$

Compared to (5.110) one concludes $\beta = -x\nu$ and then the prediction for the magnetization near the critical point for a finite system is:

$$m_0(T, L) = \xi^{-\beta/\nu} \tilde{m}(\xi/L) = L^{-\beta/\nu} \bar{m}[(1 - T/T_c)L^{1/\nu}]. \quad (5.114)$$

The typical way of checking this scaling behavior is to plot $m_0(T, L)L^{\beta/\nu}$ *vs* the rescaled variable $(1 - T/T_c)L^{1/\nu}$, see figure 5.9. One can use as well the corresponding scaling relations for the specific heat and the susceptibility:

$$c(T, L) = L^{\alpha/\nu} \bar{c}[(1 - T/T_c)L^{1/\nu}], \quad (5.115)$$

$$\chi_T(T, L) = L^{\gamma/\nu} \bar{\chi}[(1 - T/T_c)L^{1/\nu}]. \quad (5.116)$$

The scaling relation for χ_T has been checked in figure 5.10. We have to say that the quality of the scaling observed in these figures 5.9-5.10 is not representative of many simulations in other systems. When plotting the figures with the rescaled data we

have used the **known** values for the critical temperature T_c and the critical exponents ν, β and γ for the 2-d Ising model. If, as it is usually the case, the critical temperature and the critical exponents are not known, but its knowledge are the ultimate goal of our simulation, this procedure implies a three-parameter fit which is rather difficult to do in practice. The fitting has another important problem, namely, that the scaling relations, such as (5.114), (5.115), (5.116) and others are only valid *asymptotically* for large L and close enough to T_c. What is meant by "large L" and "sufficiently close to T_c" is something that can not be asserted before we do the simulations. If we add on top of all that the large errors that occur near the critical region, it is not strange that some disputes appear continuously about the correct values of critical exponents for different models of interest in the literature.

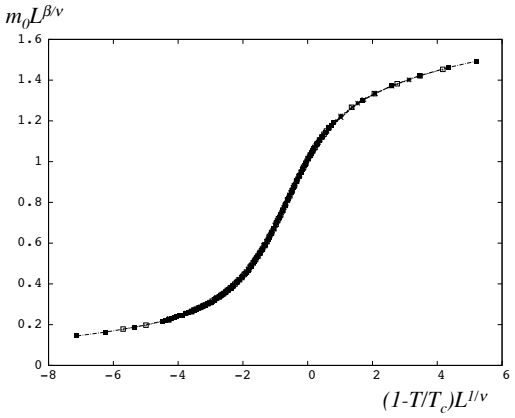

Figure 5.9 Check of the scaling relation with system size of the magnetization. Data for $L = 40, 80, 120, 160, 200$ have been collapsed onto the same curve using the rescaling of axes as indicated.

Of particular interest for the analysis of the data is the finite-size scaling behavior of the fourth-order cumulant defined as the ratio of moments of the magnetization $\mathtt{m}$ defined:

$$U_4(T, L) \equiv 1 - \frac{\langle \mathtt{m}^4 \rangle}{3\langle \mathtt{m}^2 \rangle^2} = \bar{U}_4[(1 - T/T_c)L^{1/\nu}]. \tag{5.117}$$

As $U_4(T_c, L) = \bar{U}_4(0)$, a constant, the critical temperature T_c can be then determined as the common intersection of the $U_4(T, L)$ curves for different values of L, see figure 5.11. Once T_c has been determined by this procedure, we can use the fits to determine the values of the critical exponent ν and use those values in the analysis of the curves for the magnetization, susceptibility, etc. However, be aware that these are not easy fits and it is essential to have a precise determination of the errors, including those due to the large autocorrelation times at the critical point.

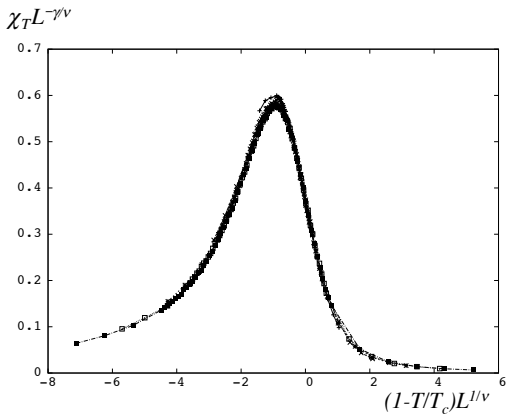

Figure 5.10 Check of the scaling relation with system size of the susceptibility. Data for $L = 40, 80, 120, 160, 200$ have been collapsed onto the same curve using the rescaling of axes as indicated.

Increase of fluctuations The variance $\sigma[m]$ measures the fluctuations of m. As we have discussed, the fluctuations increase at the critical region of a phase transition. This result can be obtained from the relation between the fluctuations and some response functions, such as (5.49) and (5.51).

Near a critical point, the isotherm becomes flat and hence some of these derivatives become formally infinite and have associated, therefore, fluctuations of a very large amount. These increase of fluctuations at the critical point is the origin of the so-called critical opalescence that can be observed with the naked eye and that constitutes one of the trademarks of critical points. But from the numerical point of view, this is a disaster. If the fluctuations, as measure by $\sigma[m]$ increase, so does the error $\epsilon[m]$. The theory of finite-size scaling that we have briefly reviewed tells us that fluctuations of the order parameter behave at the critical region as dictated by (5.116). Therefore, exactly at $T = T_c$, it is $\chi_T(T_c, L) = L^{\gamma/\nu}\bar{\chi}(0)$. If we know the value of T_c we can fit this expression to obtain γ/ν. There is another way in which we can proceed. As seen in figure 5.8 the susceptibility $\chi_T(T, L)$ develops a maximum at a location $T_c(L)$. By finding the maximum of (5.116) we can prove that the location of this maximum varies as

$$T_c(L) \quad = \quad T_c + a_1 L^{-1/\nu}, \tag{5.118}$$

being a a constant. From here we deduce that a plot of $T_c(L)$ *vs* $L^{-1/\nu}$ must yield a straight line whose interception at the origin is precisely T_c. This can provide a confirmation of the values of T_c and ν obtained by other fits, such as the ones coming from the fourth-order cumulant.

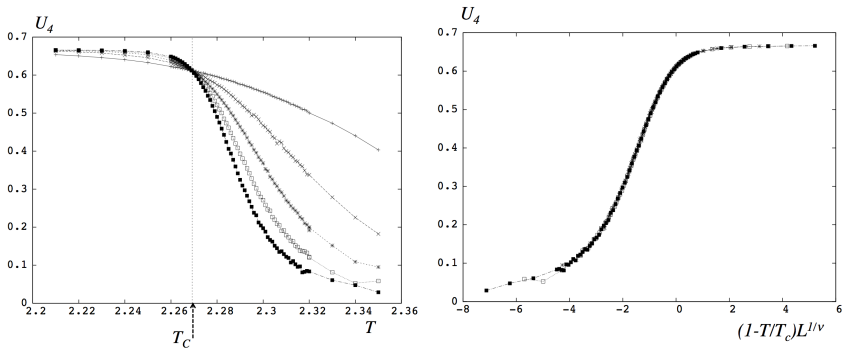

Figure 5.11 Behavior of the ratio of moments U_4 defined in (5.117) coming from numerical simulations of the 2-d square Ising model for system sizes $L = 40, 80, 120, 160, 200$. The common crossing of the curves for different values of L indicates the critical temperature T_c (left panel). In the right panel, the data have been collapsed onto the same curve using the rescaling of the horizontal axis as indicated.

Critical slowing down As well as the fluctuations, the correlation time $\tau_{\mathtt{m}}$ also diverges at the critical point! It is known that the *real* dynamics of a physical system slows down at the critical region (again, a fact that can be observed experimentally). However, we are not using here the real (say Hamiltonian) dynamics, but we have introduced (within some arbitrariness) a convenient Markov chain that allows us to generate representative configurations at equilibrium. Does this stochastic dynamics also suffer from critical slowing down? The answer is yes, as shown by the numerical data and supported by a simple calculation. We had derived in (4.80) an exact expression for the value of the correlation function $\rho_G(1)$ after one elementary proposal-acceptance step. We combine this with the approximate expression for the correlation time (**??**) to obtain

$$\tau_{\mathtt{m}} \approx \frac{2\sigma^2[\mathtt{m}]}{\int dx\, dy\, h(x|y) g(x|y) [\mathtt{m}(x) - \mathtt{m}(y)]^2 f_{\hat{\mathbf{x}}}(y)}. \tag{5.119}$$

Consider, for the sake of clarity, the Metropolis algorithm for the Ising model and we want to consider the correlation time of the magnetization (5.46). We know that the proposal for change $x \to y$ is to reverse the sign of a single spin. This yields a modification of $\mathtt{m}(x) - \mathtt{m}(x) = \frac{\pm 2}{N}$. After squaring, it can be taken out of the integral and we are left with $\int dx\, dy\, h(x|y) g(x|y) f_{\hat{\mathbf{x}}}(y)$ which is nothing but the average acceptance probability ϵ. We get finally

$$N^{-1}\tau_{\mathtt{m}} = \frac{N\sigma^2[m]}{2\epsilon} = \frac{kT\chi_T}{2\epsilon}. \tag{5.120}$$

$N^{-1}\tau_{\mathtt{m}} \equiv \tau_{\mathtt{m}}^{\mathrm{MCS}}$ is the autocorrelation time in units of updates per spin, or MCS. As the average acceptance probability does not decrease near zero at the critical point, it follows that, within this approximation, the critical behavior of the correlation time $\tau_{\mathtt{m}}^{\mathrm{MCS}}$, is that of the magnetic susceptibility χ_T. As this diverges at the critical point,

it turns out that $\tau_{\mathtt{m}}^{\mathrm{MCS}}$ diverges as well. As explained in the previous section, for a finite system of linear side L, it can not really diverge to infinity, but it grows as a power of the system side

$$\tau_{\mathtt{m}}^{\mathrm{MCS}}(T = T_c) \sim L^z. \tag{5.121}$$

The simple expression (5.120) we have derived here would tell us that $z = \gamma/\nu$, the exponent giving the divergence of the susceptibility with system size. Intensive numerical simulations using the Metropolis algorithm suggest that the dynamical critical exponent takes a value $z \approx 2$ for the 2d Ising model. As the fluctuations increase at the critical point as $\chi_T \sim L^{\gamma/\nu}$, it turns out that the error (5.109) at the critical point increases with system size as $\epsilon[m] \sim L^{(z+\gamma/\nu)/2}$ with $(z + \gamma/\nu)/2 \approx 1.9$. Roughly speaking, for the 2d Ising model, the errors in the Metropolis algorithm near the critical increase by a factor of 4 every time we double the system linear side, if keeping constant the number of measurements M.

While the divergence of the fluctuations (e.g. susceptibility) near a critical point is intrinsic to the model under study, the divergence of $\tau_{\mathtt{m}}$ and other correlation times is a property of the numerical method we use to generate the representative configurations. Therefore, some Monte Carlo schemes have a smaller correlation time than others. It is then important to compare the different possibilities available and to tune up the parameters of the numerical method, whenever possible, in order to obtain the smallest correlation time. As far as the Ising model is concerned, the best algorithms use collective updates in which many spin are changed at once. In this way it is possible to reduce the critical exponent z to a value close to 0. This, however, require an exquisite care in the choice of the proposal and acceptance probabilities. They will be discussed briefly in Appendix 15.

Thermalization We will very brief here. We just want to remind the reader that the Boltzmann factor is sampled only asymptotically by the Monte Carlo algorithm. Therefore, we need to make sure that we are in the regime where all generated configurations follow the distribution (5.3). We have already discussed that one way of checking that this is the case is to check the result (5.24). More sophisticated tests include computing the non-linear relaxation function and checking that it has decayed to a value close to 0. We refer the reader to section 4.7 for a discussion of this point.

Further reading

Concerning the applications to lattice Ising-type models, the books by M.E.J. Newman and G.T. Barkema [**?**] and by D.P. Landau and K. Binder [**?**] give further details about data analysis using finite-size scaling techniques. The second book also offers a good summary of techniques and results for off-lattice continuous models such as Lennard-Jones and hard-spheres. This topic is also covered in the classic book by Allen and Tildesley [**?**].

Exercises

1) Consider a hard-sphere model including a gravitational potential $-mgz$ acting on each sphere. Use the Monte Carlo algorithm to compute the density profile $\rho(z)$ as a function of height.
2) Prove that, in the Ising model, if we select randomly the spin to be updated, then the proportion of nodes that are not selected after N trials tends, for large N, to the value $1/e \approx 37\%$.
3) Compute numerically the autocorrelation times for the magnetization and the energy for the two-dimensional Ising model using Metropolis algorithm, as a function of temperature T for systems of size $L \times L$ with $L = 20, 40, 60, 80$. Compare the result given by the analysis of the temporal series (using, for example, the method given in Appendix 14) with the approximate value 4.26.
4) Run the Monte Carlo algorithm for the Ising model at exactly the critical temperature $T_c = 1/\log\sqrt{1+\sqrt{2}} \approx 2.2691853\ldots$ and compute the magnetization, the susceptibility and the correlation time of the magnetization for a square lattice of systems size $L \times L$ with $L = 20, 40, 80, 160, 320$ (larger if your computer allows you to do so). Check that all these quantities behave proportionally to a power of the system size L and compute in each case the exponent of the power-law. Do not forget to perform a proper analysis and give meaningful error bars for each exponent.
5) Modify the program for the 2-d Ising model including the so-called Moore neighborhood, in which each spin interacts with its 4 nearest neighbors and the 4 next-nearest.
6) Run the heat-bath algorithm for the Ising model in the 3d cubic lattice. Use the finite-size scaling techniques analysis of the fourth-order cumulant and determine the critical temperature (best estimate in the literature $T_c \approx 4.5115$).
7) In the so-called q-state Potts model, in every site of the lattice the variable s_i can take any integer value between 1 and q (the case $q = 2$ is isomorphic to the Ising model). The interaction energy between two nearest neighbors s_i, s_j is equal to 0 if $s_i \neq s_j$ and equal to -1 if $s_i = s_j$. If we want to implement the Metropolis algorithm, what is a reasonable proposal probability probability in this model? Which is the corresponding acceptance probability? Run the program for $q = 3$ and compute the order parameter, defined as

$$\mathtt{m} = \sqrt{\frac{1}{q-1}\sum_{k>\ell}(x_k - x_\ell)^2},$$

being $x_k, k = 1, \ldots, q$ the proportion of sites occupied by a value $s_i = k$. Determine the critical temperature.
8) An elegant way of implementing a Monte Carlo method for the Φ^4 model is that of Bruce[8]. In his approach, the proposal $g_{\tilde{\mathbf{x}}}(x|y) \equiv g_{\tilde{\mathbf{x}}}(\phi'_i)$ is also independent of the old configuration y (in the same vein than heat-bath), but $g_{\tilde{\mathbf{x}}}(\phi'_i)$ is chosen to be the sum of two Gaussians which best approximate the actual local distribution

of the field ϕ'_i:

$$g_{\tilde{\mathbf{x}}}(\phi'_i) = \frac{1}{2}\left[\frac{1}{\sigma_1\sqrt{2\pi}}\mathrm{e}^{-\frac{(\phi'_i-\mu_1)^2}{2\sigma_1^2}} + \frac{1}{\sigma_2\sqrt{2\pi}}\mathrm{e}^{-\frac{(\phi'_i-\mu_2)^2}{2\sigma_2^2}}\right],$$

and $\sigma_1, \sigma_2, \mu_1, \mu_2$ are determined self-consistently during the simulation. Use the detailed balance condition to determine a possible acceptance probability for this method.

9) Prove equation 5.118 for the location of the maximum of the susceptibility according to the theory of finite-size scaling.

10) Implement a simple Metropolis algorithm for the 2d square lattice Φ^4 model in which the proposal changes the scalar variable $\phi_i \to \phi'_i \in (\phi_i - \Delta, \phi_i + \Delta)$. Run the program at the approximate value of the critical point ($u = 1, b \approx -2.735$) and determine the optimal value for Δ as the one that minimizes the correlation time of the magnetization.