15

Computational methods for lattice field theories

15.1 Introduction

Classical field theory enables us to calculate the behaviour of fields within the framework of classical mechanics. Examples of fields are elastic strings and sheets, and the electromagnetic field. Quantum field theory is an extension of ordinary quantum mechanics which not only describes extended media such as string and sheets, but which is also supposed to describe elementary particles. Furthermore, ordinary quantum many-particle systems in the grand canonical ensemble can be formulated as quantum field theories. Finally, classical statistical mechanics can be considered as a field theory, in particular when the classical statistical model is formulated on a lattice, such as the Ising model on a square lattice, discussed in Chapter 7.

In this chapter we shall describe various computational techniques which are used to extract numerical data from field theories. *Renormalisation* is a procedure without which field theories cannot be formulated consistently in continuous space-time. In computational physics, we formulate field theories usually on a lattice, thereby avoiding the problems inherent to a continuum formulation. Nevertheless, understanding the renormalisation concept is essential in lattice field theories in order to make the link to the real world. In particular, we want to make predictions about physical quantities (particle masses, interaction constants) which are independent of the lattice structure, and this is precisely where we need the renormalisation concept.

Quantum field theory is difficult. It does not belong to the standard repertoire of every physicist. We try to explain the main concepts and ideas before entering into computational details, but unfortunately we cannot give proofs and derivations, as a thorough introduction to the field would require a book on its own. For details, the reader is referred to refs. 1–5. In the next section we shall briefly describe what quantum field theory is, and present several examples. In the following section,

the procedure of numerical quantum field theory will be described in the context of renormalisation theory. Then we shall describe several algorithms for simulating field theory, and in particular methods for reducing critical slowing down, a major problem in numerical field theory computations. Finally, we shall consider some applications in quantum electrodynamics (QED) and quantum chromodynamics (QCD).

15.2 Quantum field theory

To understand quantum field theory, it is essential to be accustomed to the path-integral formalism, see Section 12.4, so let us recall this concept briefly.

Consider a single particle in one dimension. The particle can move along the x-axis and its trajectory can be visualised in 1+1 dimensional space–time. Fixing initial and final positions and time to $(t_\mathrm{i}, x_\mathrm{i})$, $(t_\mathrm{f}, x_\mathrm{f})$ respectively, there is (in general) one particular curve in the (t, x)-plane, the *classical trajectory*, for which the action S is stationary. The action is given by

$$S(x_\mathrm{i}, x_\mathrm{f}; t_\mathrm{i}, t_\mathrm{f}) = \int_{t_\mathrm{i}}^{t_\mathrm{f}} d\,t\, L(x, \dot{x}, t) \tag{15.1}$$

where $L(x, \dot{x}, t)$ is the Lagrangian. In quantum mechanics, nonstationary paths are allowed too, and the probability to go from an initial position x_i to a final position x_f is given by

$$\int [\mathscr{D}x(t)] e^{-iS/\hbar} = \left\langle x_\mathrm{f} \middle| e^{-i(t_\mathrm{i}-t_\mathrm{f})H/\hbar} \middle| x_\mathrm{i} \right\rangle, \tag{15.2}$$

where H is the Hamiltonian of the system. The integral $\int [\mathscr{D}x(t)]$ is over all possible paths with fixed initial and final values x_i and x_f respectively. If we send Planck's constant $\hbar$ to zero, the significant contributions to the path integral will be more and more concentrated near the stationary paths, and the stationary path with the lowest action is the only one which survives when $\hbar = 0$.

Now consider a field. The simplest example of a field is a one-dimensional string, which we shall consider as a chain of particles with mass m, connected by springs such that in equilibrium the chain is equidistant with spacing a. The particles can move along the chain – the displacement of particle n with respect to the equilibrium position is called ϕ_n. Fixed, free, or periodic boundary conditions can be imposed. The chain is described by the action

$$S = \frac{1}{2} \int_{t_\mathrm{i}}^{t_\mathrm{f}} \left\{ \sum_n \frac{1}{2} m \dot{\phi}_n^2(t) - A \left[\frac{\phi_{n+1}(t) - \phi_n(t)}{a} \right]^2 \right\} d\,t. \tag{15.3}$$

A is a constant, and some special conditions are needed for the boundaries. In a quantum mechanical description we again use the path integral, which gives us

the probability density for the chain to go from an initial configuration $\Phi^{\mathrm{i}} = \{\phi_n^{(\mathrm{i})}\}$ at time t_{i} to a final configuration $\Phi^{\mathrm{f}} = \{\phi_n^{(\mathrm{f})}\}$ at time t_{f} (note that $\Phi^{(\mathrm{i,f})}$ denotes a complete chain configuration):

$$\int [\mathscr{D}\Phi(t)] e^{-iS/\hbar} \tag{15.4}$$

where the path integral is over all possible configurations of the field Φ in the course of time, with fixed initial and final configurations.

We want to formulate this problem now in continuum space. To this end we replace the discrete index n by a continuous index x_1, and we replace the interaction term occuring in the summand by the continuous derivative:

$$S = \frac{1}{2}\int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} dt \int dx_1 \left\{ m\dot{\phi}(t,x_1)^2 - A\left[\frac{\partial\phi(t,x_1)}{\partial x_1}\right]^2 \right\}. \tag{15.5}$$

The field $\phi(t,x_1)$ can be thought of as a sheet whose shape is given as a height $\phi(t,x_1)$ above the 1+1 dimensional space–time plane. In the path integral, we must sum over all possible shapes of the sheet, weighted by the factor $e^{iS/\hbar}$. The field can be rescaled at will, as it is integrated over in the path integral (this rescaling results in an overall prefactor), and the time and space units can be defined such as to give the time derivative a prefactor $1/c$ with respect to the spatial derivative (c is the speed of light), and we obtain

$$S = \int d^2x \, \frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x), \tag{15.6}$$

where we have used x to denote the combined space–time coordinate $x \equiv (t,x_1) \equiv (x_0,x_1)$. From now on, we put $c \equiv \hbar \equiv 1$ and we use the Einstein-summation convention according to which repeated indices are summed over. The partial space–time derivatives $\partial^\mu, \partial_\mu$ are denoted by:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \quad , \quad \partial^\mu = \frac{\partial}{\partial x_\mu}. \tag{15.7}$$

Furthermore we use the Minkowski metric:

$$a_\mu a^\mu = a_0^2 - \mathbf{a}^2. \tag{15.8}$$

The fact that we choose $c \equiv \hbar \equiv 1$ leaves only one dimension for distances in space–time, and masses and energies. The dimension of inverse distance is equal to the energy dimension, which is in turn equal to the mass dimension.

Using partial integration, we can reformulate the action as

$$S = -\int d^2x \, \frac{1}{2}\phi(x)\partial_\mu\partial^\mu\phi(x), \tag{15.9}$$

where we have assumed periodic boundary conditions in space and time (or vanishing fields at the integral boundaries, which are located at infinity) to neglect integrated terms.

If, apart from a coupling to its neighbours, each particle had also been coupled to an external harmonic potential $m^2\phi^2/2$, we would have obtained

$$S = -\int d^2x \, \frac{1}{2}\phi(x)(\partial_\mu \partial^\mu + m^2)\phi(x). \tag{15.10}$$

Note that the Euler-Lagrange equations for the field are

$$(\partial^\mu \partial_\mu + m^2)\phi(x) = 0; \tag{15.11}$$

which is recognised as the Klein-Gordon equation, the straightforward covariant generalisation of the Schrödinger equation.[†]

Quantum field theory is often used as a theory for describing particles. The derivation above started from a chain of particles, but these particles are merely used to formulate our quantum field theory, and they should not be confused with the physical particles which are described by the theory. The difference between the two can be understood as follows. Condensed matter physicists treat wave-like excitations of the chain (i.e. a one-dimensional 'crystal') as particles – they are called *phonons*. Note that the 'real' particles are the atoms of the crystal. In field theory, the only 'real' particles are the excitations of the field.

In fact, we can imagine that a wave-like excitation pervades our sheet, for example

$$\phi(t, x_1) = e^{ipx_1 - i\omega t} \tag{15.12}$$

(here x_1 denotes the spatial coordinate). This excitation carries a momentum p and an energy $\hbar\omega$, and it is considered as a particle. We might have various waves as a superposition running over the sheet: these correspond to as many particles.

Let us try to find the Hamiltonian corresponding to the field theory presented above (the following analysis is taken up in some detail in problems 15.2 and 15.3). We do this by returning to the discretised version of the field theory. Let us first consider the ordinary quantum description of a single particle of mass 1, moving in one dimension in a potential $mx^2/2$. The Hamiltonian of this particle is given by

$$H = \frac{p^2}{2} + \frac{m}{2}x^2 \tag{15.13}$$

with $[p, x] = -i$. In the example of the chain we have a large number of such particles, but each particle can still be considered as moving in a harmonic potential,

[†]The Klein-Gordon equation leads to important problems in ordinary quantum mechanics, such as a nonconserved probability density and an energy spectrum which is not bounded from below.

and after some calculation we find for the Hamiltonian:

$$H = \sum_n \left[\hat{\pi}_n^2 + (\hat{\phi}_n - \hat{\phi}_{n-1})^2 + m^2\hat{\phi}_n^2\right], \tag{15.14}$$

with

$$[\hat{\pi}_n, \hat{\phi}_l] = -i\delta_{nl}. \tag{15.15}$$

The hats are put above ϕ and π to emphasise that they are now operators. The Hamiltonian can be diagonalised by first Fourier transforming and then applying operator methods familiar from ordinary quantum mechanics to it. The result is[2, 6]

$$H = \frac{1}{2}\int_{-\pi}^{\pi} dk\, \omega_k \hat{a}_k^\dagger \hat{a}_k \tag{15.16}$$

where $\hat{a}_k^\dagger$ is a creation operator: it creates a Fourier mode

$$\phi_n = e^{ikn} \tag{15.17}$$

and $\hat{a}_k$ is the corresponding destruction or annihilation operator. In the ground state (the 'vacuum') there are no modes present and the annihilation operator acting on the ground state gives zero:

$$a_k\,|0\rangle = 0. \tag{15.18}$$

The Fourier modes represent energy quanta of energy $\hbar\omega$; the *number operator* $n_k = a_k^\dagger a_k$, acting on a state $|\psi\rangle$ counts the number of modes (quanta) with wave vector k, present in that state. The Hamiltonian (15.16) operator then adds all the energy quanta which are present in the state.

In fact, $\hat{a}_k$ is given in terms of the Fourier transforms of the $\hat{\phi}$ and $\hat{\pi}$ operators:

$$\hat{a}_k = \frac{1}{\sqrt{4\pi\omega_k}}\left[\omega_k\hat{\phi}_k + i\hat{\pi}_k\right], \tag{15.19}$$

analogous to the definition of creation and annihilation operators for the harmonic oscillator. The frequency ω is related to k by

$$\omega_k = \omega_{-k} = \sqrt{m^2 + 2(1-\cos k)}. \tag{15.20}$$

For small k we find the continuum limit:

$$\omega_k = \sqrt{m^2 + k^2} \tag{15.21}$$

which is the correct dispersion relation for a relativistic particle with mass m (in units where $c = \hbar = 1$).

We see that the Hamiltonian (15.16) of a particular configuration is simply given as the sum of the energies of a set of one-particle Hamiltonians (remember these particles are nothing but Fourier-mode excitations of the field): the particles do not interact. Therefore, the field theory considered so far is called *free field theory*. The eigenstates of the free field theory are

$$|k_1, \ldots, k_M\rangle = \hat{a}^\dagger_{k_1} \ldots \hat{a}^\dagger_{k_M} |0\rangle \tag{15.22}$$

for arbitrary M, which denotes the number of particles present. It is possible to have the same k_i occurring more than once in this state (with an appropriate normalisation factor): this means that there is more than one particle with the same momentum. The state $|0\rangle$ is the vacuum state: it corresponds to having no particles. The lowest energy above the vacuum energy is that corresponding to a single particle at rest ($k = 0$): the energy is equal to the mass. In Chapter 11 we have seen that for a statistical field theory the inverse of the lowest excitation energy is equal to the correlation length:

$$m \sim 1/(\xi a). \tag{15.23}$$

However, this holds for a statistical field theory where we do not have complex weight factors – these can be made real by an analytical continuation of the physics into imaginary time: $t \to it$ (see also Section 12.2.4). In that case the (continuous) action in d space–time dimensions reads

$$S = \int d^d x \, \frac{1}{2} \left[\partial_\mu \phi \partial^\mu \phi + m^2 \phi^2 \right] \tag{15.24}$$

where now

$$\partial_\mu \phi \partial^\mu \phi = (\nabla \phi)^2 + \left(\frac{\partial}{\partial t} \phi \right)^2 \tag{15.25}$$

i.e. the Minkowski metric has been replaced by the Euclidean metric. The matrix elements of the time evolution operator now read $\exp(-S)$ instead of $\exp(-S/i)$ (for $\hbar \equiv 1$). We have now a means to determine the particle mass: simply by measuring the correlation length. In the free field theory, the inverse correlation length is equal to the mass parameter m in the Lagrangian, but if we add extra terms to the Lagrangian (see below) then the inverse correlation length (or the physical mass) is no longer equal to m.

It might seem that we have been a bit light-hearted in switching to the Euclidean field theory. Obviously, expectation values of physical quantities can be related for the Minkowski and Euclidean versions by an analytic continuation. In the numerical simulations we use the Euclidean metric to extract information concerning the Hamiltonian: this operator is the same in both metrics – only the

time evolution, and hence the Lagrangian changes when going from one metric to the other. Euclidean field theory can therefore be considered merely as a trick to study the spectrum of a quantum Hamiltonian of a field theory which in reality lives in Minkowski space.

If we add another term to the Lagrangian:

$$S = \int d^d x \left\{ \frac{1}{2}\phi(x)(-\partial_\mu \partial^\mu + m^2)\phi(x) + V[\phi(x)] \right\}, \tag{15.26}$$

where V is not quadratic (in that case it would simply contribute to the mass term), then interactions between the particles are introduced. Usually one considers

$$V = g\phi^4(x)/2 \tag{15.27}$$

and the Lagrangian describes the simplest interesting field theory for interacting particles, the *scalar ϕ^4 theory*. The name 'scalar' denotes that $\phi(x)$ is not a vector. Vector theories do exist, we shall encounter examples later on. When a potential is present, the energy is no longer a sum of one-particle energies: the particles interact.

We have mentioned the probability to go from a particular initial state to another (final) state as an example of the problems studied in field theory. Our experimental knowledge on particles is based on scattering experiments. This is a particular example of such a problem: given two particles with certain initial states, what are the probabilities for different resulting reaction products, that is, which particles do we have in the end and what are their momenta? In scalar field theory we have only one type of particle present. As we have seen in the first chapter of this book, experimental information on scattering processes is usually given in terms of scattering cross sections. These scattering cross sections can be calculated – they are related to an object called *S-matrix*. The S-matrix is defined as

$$S_{\mathrm{fi}} = \lim_{\substack{t_\mathrm{i}\to-\infty \\ t_\mathrm{f}\to\infty}} \langle \psi_\mathrm{f} | U(t_\mathrm{i}, t_\mathrm{f}) | \psi_\mathrm{i} \rangle . \tag{15.28}$$

Our initial state is one with a particular set of initial momenta as in (15.22), and similarly for the final state; $U(t_\mathrm{i}, t_\mathrm{f})$ is the time evolution operator,[†] and the states $\psi_{\mathrm{i,f}}$ usually contain a well-defined number of free particles with well-defined momenta (or positions, depending on the representation).

Scattering cross sections are expressed directly in terms of the S-matrix, and the latter is related to the *Green functions* of the theory by the so-called Lehmann–Symanzik–Zimmermann relation, which can be found for example in Ref. 2. These

[†]More precisely, the time evolution operator is that for a theory with an interaction switched on at a time much later than t_i and switched off again at a time long before t_f.

Green functions depend on a set of positions $x_1, \ldots, x_n$ and are given by

$$\mathcal{G}(x_1, \ldots, x_n) = \int [\mathcal{D}\phi]\phi(x_1)\cdots\phi(x_n)e^{-S/\hbar} / \int [\mathcal{D}\phi]e^{-S/\hbar}. \tag{15.29}$$

Note that x_i is a space–time vector, the subscripts do not denote space–time components. The scattering cross sections are evaluated in the Euclidean metric – the Minkowskian quantities are obtained by analytical continuation: $t \to it$.

As the initial and final states in a scattering experiment are usually given by the particle momenta, we need the Fourier transform of the Green's function, defined as

$$\mathcal{G}(k_1, \ldots, k_n)\delta(k_1 + \cdots + k_n)(2\pi)^d = \int d^d x_1 \ldots d^d x_n e^{ik_1\cdot x_1 + \cdots + ik_n \cdot x_n}\mathcal{G}(x_1, \ldots, x_n). \tag{15.30}$$

The d-dimensional delta-function reflects the energy-momentum conservation of the scattering process, which is related to the space–time translation invariance of the Green's function.

For the free field theory it is found that

$$\mathcal{G}(k, -k) = \frac{1}{k^2 + m^2} \tag{15.31}$$

which leads to the real-space form:

$$\mathcal{G}(x - x') = \frac{e^{-|x-x'|m}}{|x - x'|^\eta}; \qquad \text{large } |x - x'|, \tag{15.32}$$

with $\eta = (d-1)/2$. We see that the Green function has a finite correlation length $\xi = 1/m$. Higher-order correlation functions for the free field theory can be calculated using *Wick's theorem*: correlation functions with an odd number of ϕ-fields vanish, but if they contain an even number of fields, they can be written as a symmetric sum over products of pair-correlation functions, for example

$$G(x_1, x_2, x_3, x_4) = \langle \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle = \langle \phi(x_1)\phi(x_2)\rangle\langle \phi(x_3)\phi(x_4)\rangle + \langle \phi(x_1)\phi(x_3)\rangle\langle \phi(x_2)\phi(x_4)\rangle + \langle \phi(x_1)\phi(x_4)\rangle\langle \phi(x_2)\phi(x_3)\rangle. \tag{15.33}$$

In fact, it is well known that for stochastic variables with a Gaussian distribution, all higher moments can be formulated similarly in terms of the second moment – Wick's theorem is a generalisation of this result.

15.3 Interacting fields and renormalisation

The free field theory can be solved analytically: all the Green functions can be given in closed form. This is no longer the case when we are dealing with interacting fields. If we add for example a term $g\phi^4$ to the free field Lagrangian, the only way to proceed analytically is by performing a perturbative analysis in the coupling constant g. It turns out that this gives rise to rather difficult problems. The terms in the perturbation series involve integrals over some momentum coordinates, and these integrals diverge! Obviously our predictions for physical quantities must be finite numbers, so we seem to be in serious trouble. Since this occurs in most quantum field theories as soon as we introduce interactions, it is a fundamental problem which needs to be faced.

To get a handle on the divergences, one starts by controlling them in some suitable fashion. One way to do this is by cutting off the momentum integrations at some large but finite value Λ. This renders all the integrals occurring in the perturbation series finite, but physical quantities depend on the – arbitrary – cut-off Λ which is still unaccepTable. Another way to remove the divergences is by formulating the theory on a discrete lattice. This is of course similar to cutting of momentum integrations, and the lattice constant a used is related to the momentum cut-off by

$$a \sim 1/\Lambda. \tag{15.34}$$

Such cut-off procedures are called *regularisations* of the field theory.

We must remove the unphysical cut-off dependence from the theory. The way to do this is to allow the coupling constant and mass constants of the theory to be cut-off dependent and then require that the cut-off dependency of the Green functions disappears.[†] There are infinitely many different Green's functions and it is not obvious that these can all be made cut-off independent by adjusting only the three quantities m, g and ϕ. Theories for which this *is* possible are called *renormalisable*. The requirement that all terms in the perturbative series are merely *finite* without a prescription for the actual values, leaves some arbitrariness in the values of field scaling, coupling constant and mass – we use experimental data to fix these quantities.

To be more specific, suppose we carry out the perturbation theory to some order. It turns out that the resulting two-point Green's function $\mathcal{G}(k,-k)$ assumes the form of the free-field correlation function (15.31) with a finite mass parameter plus some cut-off dependent terms. Removing the latter by choosing the various constants of the theory (m, g, scale of the field) in a suitable way, we are left with

$$\mathcal{G}(k,-k) = 1/\left(k^2 + m_{\mathrm{R}}^2\right) \tag{15.35}$$

[†]In addition to mass and coupling constant, the field is rescaled by some factor.

where m_R is called the 'renormalised mass' – this is the physical mass which is accessible to experiment. This is not the mass which enters in the Lagrangian and which we have made cut-off dependent – the latter is called the 'bare mass', which we shall now denote by m_B. The value of the renormalised mass m_R is not fixed by the theory, as the cut-off removal is prescribed up to a constant. We use the experimental mass to 'calibrate' our theory by fixing m_R. In a similar fashion, we use the experimental coupling constant, which is related to the four-point Green function, to fix a renormalised coupling constant g_R (see the next section).

The renormalisation procedure sounds rather weird, but it is certainly not some arbitrary *ad hoc* scheme. The aim is to find bare coupling constants and masses, such that the theory yields cut-off independent physical (renormalised) masses and couplings. Different regularisation schemes all lead to the same physics. We need as many experimental data as we have parameters of the theory to adjust, and having fixed these parameters we can predict an enormous amount of new data (in particular, all higher order Green's functions). Moreover, the requirement that the theory is renormalisable is quite restrictive – for example, only the ϕ^4 potential has this property; changing the ϕ^4 into a ϕ^6 destroys the renormalisability of the theory.

In computational physics we usually formulate the theory on a lattice. We then choose values for the bare mass and coupling constant and calculate various physical quantities in units of the lattice constant a (or its inverse). Comparison with experiment then tells us what the actual value of the lattice constant is. Therefore the procedure is somehow the reverse of that followed in ordinary renormalisation, although both are intimately related. In ordinary renormalisation theory we find the bare coupling constant and mass as a function of the cut-off from a comparison with experiment. In computational field theory we find the lattice constant as a function of the bare coupling constant from comparison with experimental data.

Let us consider an example. We take the Euclidean ϕ^4 action in dimension $d = 4$:

$$S = \frac{1}{2}\int d^4x \left\{ [\partial_\mu \phi(x)][\partial^\mu \phi(x)] + m^2\phi^2(x) + g\phi^4(x) \right\} \tag{15.36}$$

and discretise this on the lattice, with a uniform lattice constant a. Lattice points are denoted by the four-index $n = (n_0, n_1, n_2, n_3)$. A lattice point n corresponds to the physical point $x = (an_0, an_1, an_2, an_3) = an$. The discretised lattice action reads

$$S_{\text{Lattice}} = \frac{1}{2}\sum_n a^4 \left\{ \sum_{\mu=0}^{3} \left[\frac{\phi(n+e_\mu) - \phi(n)}{a} \right]^2 + m^2\phi_n^2 + g\phi^4(n) \right\}. \tag{15.37}$$

We rescale the ϕ-field, the mass and the coupling constant according to

$$\phi(n) \to \phi(n)/a; \quad m \to m/a \quad \text{and} \quad g \to g, \tag{15.38}$$

to make the lattice action independent of the lattice constant a:

$$S_{\text{Lattice}} = \frac{1}{2}\sum_n \left\{ \sum_{\mu=0}^{3} \left[\phi(n+e_\mu) - \phi(n)\right]^2 + m^2\phi_n^2 + g\phi^4(n) \right\}. \tag{15.39}$$

Now we perform an MC or another type of simulation for particular values of m and g. We can then 'measure' the correlation length in the simulation. This then should be the inverse of the experimental mass, measured in units of the lattice constant a. Suppose we know this mass from experiment, then we can infer what the lattice constant is in real physical units.

Life is however not as simple as the procedure we have sketched suggests. The problem is that in the generic case, the correlation length is quite small in units of lattice constants. However, a lattice discretisation is only allowed if the lattice constant is *much smaller* than the typical length scale of the physical problem. Therefore, the correlation length should be an order of magnitude larger than the lattice constant. Only close to a critical point does the correlation length assume values much larger than the lattice constant. This means that our parameters m and g should be chosen close to a critical point. The ϕ^4 theory in $d = 4$ dimensions has one critical line,[7] passing through the point $m = g = 0$, the massless free field case. Therefore, m and g should be chosen very close to this critical line in order for the lattice representation to be justifiable.

As the experimental mass of a particle is a fixed number, varying the lattice constant a forces us to vary g and m in such a way that the correlation length remains finite. Unfortunately this renders the use of finite size scaling techniques impossible: the system size L must always be larger than the correlation length: $a \ll \xi < L$. †

The fact that the lattice field theory is always close to a critical point implies that we will suffer from *critical slowing down*. Consider a Monte Carlo (MC) simulation of the field theory. We change the field at one lattice point at a time. At very high temperature, the field values at neighbouring sites are more or less independent, so after having performed as many MC attempts as there are lattice sites (one MCS), we have obtained a configuration which is more or less statistically independent from the previous one. If the temperature is close to the

† In the case where physical particles are massless, so that the correlation length diverges, finite size scaling can be applied. Finite size scaling applications in massive particle field theories have however been proposed; see Ref. 8.

critical temperature, however, fields at neighbouring sites are strongly correlated, and if we attempt to change the field at a particular site, the coupling to its neighbours will hardly allow a significant change with respect to its previous value at that site. However, in order to arrive at a statistically independent configuration, we need to change the field over a volume of linear size equal to the correlation length. If that length is large, it will obviously take a very long time to change the whole region, so this problem gets worse when approaching the critical point. Critical slowing down is described by a dynamic critical exponent z which describes the divergence of the decay time τ of the dynamic correlation function [see Chapter 7, Eq. (7.73)]:

$$\tau = \xi^z, \tag{15.40}$$

where ξ is the correlation length of the system.

In recent years, much research has aimed at finding simulation methods for reducing the critical time relaxation exponent. In the following section we shall describe a few straightforward methods developed for simulating quantum field theories, using the ϕ^4 scalar field theory in two dimensions as a testing model. In Section 15.5 we shall focus on methods aiming at reducing critical slowing down. We shall then also discuss methods devised for the Ising model and for a two-dimensional model with continuous degrees of freedom.

In the final sections of this chapter, simulation methods for gauge field theories (QED, QCD) will be discussed.

15.4 Algorithms for lattice field theories

We start by reviewing the scalar Euclidean ϕ^4 field theory in d dimensions in more detail. The continuum action is

$$S_{\mathrm{E}} = \frac{1}{2}\int d^d x \left[\partial_\mu \phi(x)\partial^\mu \phi(x) + m^2\phi^2(x) + g\phi^4(x)\right] \tag{15.41}$$

(the subscript E stands for Euclidean). For $g = 0$, we have the free field theory, describing noninteracting spinless bosons. Performing a partial integration using Green's first identity, and assuming vanishing fields at infinity, we can rewrite the action as

$$S_{\mathrm{E}} = \frac{1}{2}\int d^d x \left[-\phi(x)\partial_\mu\partial^\mu \phi(x) + m^2\phi^2(x) + g\phi^4(x)\right]. \tag{15.42}$$

The scalar field theory can be formulated on a lattice by replacing derivatives by finite differences. We can eliminate the dependence of the lattice action on the lattice constant by rescaling the field, mass and coupling constant according to

$$\hat{\phi}_n = a^{d/2-1}\phi(an); \quad \hat{m} = am; \quad \hat{g} = a^{4-d}g. \tag{15.43}$$

For the four-dimensional case, $d = 4$, we have given these relations already in the previous section. Later we shall concentrate on the two-dimensional case, for which the field ϕ is dimensionless. In terms of the rescaled quantities, the lattice action reads:

$$S_{\mathrm{E}}^{\mathrm{Lattice}} = \frac{1}{2}\sum_n \left[-\sum_\mu \hat{\phi}_n \hat{\phi}_{n+\mu} + (2d + \hat{m}^2)\hat{\phi}_n^2 + \hat{g}\hat{\phi}_n^4 \right]. \tag{15.44}$$

The arguments n are vectors in d dimensions with integer coefficients and the sum over μ is over all positive and negative Cartesian directions. The action (15.44) is the form which we shall use throughout this section and it will henceforth be denoted as S. From now on we shall omit the carets from the quantities occurring in the action (15.44). As we shall simulate the field theory in the computer, we must make the lattice finite – the linear size is L.

We now describe the analytical solution of this lattice field theory for the case $g = 0$ (free field theory). The free field theory action is quadratic – it can be written in the form

$$S_{\mathrm{E}} = \frac{1}{2}\sum_{nl} \phi_n K_{nl} \phi_l, \tag{15.45}$$

where

$$K_{nl} = (2d + m^2)\delta_{nl} - \sum_\mu \delta_{n,l+\mu}. \tag{15.46}$$

Defining Fourier-transformed fields as usual:

$$\phi_k = \sum_n \phi_n e^{ik\cdot n}; \tag{15.47a}$$

$$\phi_n = \frac{1}{L^d}\sum_k \phi_k e^{-ik\cdot n}, \tag{15.47b}$$

where n and l run from 0 to $L-1$, periodic boundary conditions are assumed, and the components of k assume the values $2m\pi/L$, $m = 0, \ldots, L-1$. Then we can rewrite the free-field action as

$$S_{\mathrm{E}} = \frac{1}{2L^{2d}}\sum_k \phi_k K_{k,-k} \phi_{-k}, \tag{15.48}$$

as $K_{k,-k}$ are the only nonzero elements of the Fourier transform $K_{k,k'}$:

$$K_{k,k'} = L^d \left[-\sum_\mu 2\cos(k_\mu) + (2d + m^2) \right] \delta_{k,-k'} = L^d \left[4\sum_\mu \sin^2 \frac{k_\mu}{2} + m^2 \right] \delta_{k,-k'} \tag{15.49}$$

where the sum is now only over the positive μ directions; k_μ is the μ-component of the Fourier wavevector k.

The partition function

$$Z = \int [\mathscr{D}\phi_k] \exp\left(-\frac{1}{2L^{2d}} \sum_k \phi_k K_{k,-k} \phi_{-k}\right) = \int [\mathscr{D}\phi_k] \exp\left(-\frac{1}{2L^{2d}} \sum_k |\phi_k|^2 K_{k,-k}\right) \quad (15.50)$$

(up to a normalisation factor) is now a product of simple Gaussian integrals, with the result ($N = L^d$):

$$Z = (2\pi N^2)^{N/2} / \prod_k \sqrt{K_{k,-k}} = (2\pi N^2)^{N/2} / \sqrt{\det K} = (2\pi N^2)^{N/2} \sqrt{\det(K^{-1})}. \quad (15.51)$$

The partition function appears as usual in the denominator of expressions for expectation values. We can calculate for example the two-point correlation or Green's function $\langle \phi_n \phi_l \rangle$. The Fourier transform of this correlation function can be found quite easily:

$$\langle \phi_n \phi_l \rangle = \frac{1}{L^{2d}} \sum_{k,k'} \langle \phi_k \phi_{k'} \rangle \, e^{ik\cdot n} e^{ik'\cdot l}; \quad (15.52a)$$

$$\langle \phi_k \phi_{k'} \rangle = \frac{L^{2d}}{K_{k,-k}} \delta_{k,-k'}. \quad (15.52b)$$

Taking the small k limit in (15.49) and (15.52) leads to the form (15.31), as it should be. Taking $k = 0$, we find

$$\langle \phi^2_{k=0} \rangle = \left\langle \left(\sum_n \phi_n \right)^2 \right\rangle = L^d \zeta / m_{\mathrm{R}}^2, \quad (15.53)$$

where the factor ζ in the right hand side represents the square of the scaling factor of the field – from (15.43), $\zeta = a^{d-2}$. This equation enables us therefore to determine ζ / m_{R} in a simulation: simply by calculating the average value of $\langle \Phi^2 \rangle$, $\Phi = \sum_n \phi_n$.

We have seen that according to Wick's theorem, the correlation functions to arbitrary order for free fields can always be written as sums of products of two-point correlation functions. If we switch on the ϕ^4 interaction, we will note deviations from this Gaussian behaviour to all higher orders. Renormalisation ideas suggest that it should be possible to express all higher order correlation functions in terms of second and fourth order correlation functions, if the arguments of the Green's function are not too close (that is much farther apart than the cut-off a). The second

order Green's functions are still described by the free field form (15.52), but with m in the kernel $K_{k,k'}$ being replaced by a *renormalised mass*, m_{R}. The deviations from the Gaussian behaviour manifest themselves in fourth and higher order correlation functions – therefore a natural definition of the renormalised coupling constant g_{R} is

$$g_{\mathrm{R}} = \frac{\langle \Phi^4 \rangle - 3 \langle \Phi^2 \rangle^2}{\langle \Phi^2 \rangle^2} \tag{15.54}$$

where $\Phi = \sum_n \phi_n$.[†] Equations (15.53) and (15.54) are used below to measure the renormalised mass and coupling constant in a simulation.

15.4.1 Monte Carlo methods

The problem of calculating expectation values for the interacting scalar field theory is exactly equivalent to the problem of finding expectation values of a statistical field theory. Therefore we can apply the standard Monte Carlo (MC) algorithms of Chapter 10 straightforwardly in order to sample field configurations with Boltzmann weight $\exp(-S[\phi])$. Starting point is the action (15.44). An obvious method is the Metropolis MC algorithm, in which lattice sites are visited at random or in lexicographic order, and at the selected site a change in the field is attempted by some random amount. The change in the field is taken from a random number generator either uniformly within some interval or according to a Gaussian distribution (with a suitable width). Then we calculate the change in the action due to this change in the field. The trial value of the field is then accepted as the field value in the next step according to the probability

$$P_{\mathrm{Accept}} = e^{-S[\phi_{\mathrm{new}}]+S[\phi_{\mathrm{old}}]} \tag{15.55}$$

where the exponent on the right hand side is the difference between the action of the new and old field at the selected site, keeping the field at the remaining sites fixed. If $P_{\mathrm{Accept}} > 1$, then the new configuration is accepted.

In Chapter 10 we have already encountered another method which is more efficient as it reduces correlations between subsequent configurations: the heat bath algorithm. In this algorithm, the trial value of the field is chosen independently of the previous value. Let us call $W_\phi[\phi_n]$ the Boltzmann factor $e^{-S[\phi]}$ for a field which is fixed everywhere except at the site n. We generate a new field value at site n according to the probability distribution $W_\phi[\phi_n]$. This is equivalent to performing infinitely many Metropolis steps at the same site n successively. The

[†]This renormalisation scheme corresponds to defining the renormalised coupling constant as the four-point one-particle irreducible (OPI) Green's function in tree approximation at momentum zero.[2, 4, 5]

new value of ϕ_n can be chosen in two ways: we can generate a trial value according to some distribution $\rho(\phi_n)$ and accept this value with probability proportional to $W_\phi[\phi_n]/\rho(\phi_n)$, or we can directly generate the new value with the required probability $W_\phi[\phi_n]$. The Gaussian free field model will serve to illustrate the last method.

Consider the action (15.44). If we vary ϕ_n, and keep all the remaining field values fixed, we see that the minimum of the action occurs for $\bar{\phi}_n = \sum_\mu \phi_{n+\mu}/(2d + m^2)$, where the sum is over *all* neighbouring points, i.e. for both positive and negative directions. The Boltzmann factor $W_\phi[\phi_n]$ as a function of ϕ_n for all remaining field values fixed is then a Gaussian centred around $\bar{\phi}_n$ and with a width $1/\sqrt{2d + m^2}$. Therefore, we generate a Gaussian random number r with a variance 1, and then we set the new field value according to

$$\phi_n = \bar{\phi}_n + r/\sqrt{2d + m^2}. \tag{15.56}$$

An advantage of this method is that no trial steps have to be rejected, which obviously improves the efficiency.

Unfortunately, this method is not feasible when a ϕ^4 interaction is present as we cannot generate random numbers with an $\exp(-x^4)$ distribution. Therefore we treat this term with an acceptance/rejection step as described above. This is done as follows. First we generate a 'provisional' value of the field ϕ_n according with a Gaussian distribution $\rho(\phi_n)$, according to the procedure just described. Then we accept this new field value with a probability $\exp(-g\phi_n^4/2)$. If g is not too large (and this will be the case in most of the examples given below), then the acceptance rate will still be reasonably close to 1 and not too many trial steps are rejected. If g is large, then a different procedure for generating the trial field value should be followed.[9]

There is an intimate relation between the heat bath method described here and the Gauss-Seidel method for finding the solution of the Poisson equation (see Section A.7.2.2). In the Gauss-Seidel method, the sites are visited in lexicographic order (the same can be done in the heat bath method), and ϕ_n is set equal to $\bar{\phi}_n$ without adding a Gaussian random number to it. In Section A.7.2.2 the problem of slow convergence of the numerical solution of the Poisson problem will be addressed: it turns out that the relaxation time, measured in sweeps over the entire lattice, scales as the square of the linear lattice size. The amount of computer time involved in one lattice sweep scales also linearly with the lattice volume, so the total time needed to obtain results within a certain level of accuracy scales with the volume squared. Because of this power-law scaling behaviour of the standard Poisson solvers, one might call this problem 'critical': the relaxation time scales with the system size in a way similar to a system subject to critical fluctuations.

The relation between Poisson solvers and free field theory leads us to apply clever methods for solving Poisson's equation to the problem of generating configurations with a probability density $\exp(-S[\phi])$. In Appendix A, successive over-relaxation (SOR), the use of fast Fourier transforms (FFT), and the multigrid method are mentioned, and we shall see that all of these methods have their counterpart in Monte Carlo.

Successive over-relaxation (SOR) is a method for increasing the efficiency of the Gauss-Seidel method. The idea behind this method is that if we update the sites in lexicographic order, half of the neighbours of the site being updated have already been updated and the other half are still to be treated. In SOR, a compensation is built in for the fact that half of the neighbouring sites have not yet been updated. Site n is being updated according to

$$\phi_n^{\text{new}} = \phi_n^{\text{old}} + \omega(\bar{\phi}_n - \phi_n^{\text{old}}). \tag{15.57}$$

It can be shown that the optimal value for ω is close to 2: in that case the relaxation time, which scales as L^2 (measured in lattice sweeps) in the Gauss-Seidel method is reduced to L (see Section A.7.2.2 and Ref. 10). Adler[11] has shown that the relaxation time for a Monte Carlo algorithm where a Gaussian random number is added to ϕ^{new}:

$$\phi_n^{\text{new}} \to \phi_n^{\text{new}} + \sqrt{\omega(2-\omega)}\, r / \sqrt{2d + m^2}, \tag{15.58}$$

is equal to that of the corresponding Poisson solver algorithm, that is, the relaxation time will now scale as L. We should obviously check that the SOR method still satisfies detailed balance. This is left as an exercise, see problem 15.5.

The SOR method works well for models with quadratic interactions – including a ϕ^4 term renders the method less suitable, see however Ref. 12. Fortunately, the physically more interesting gauge theories which will be discussed later in this chapter are quadratic. A problem with this method is that the optimal value of the over-relaxation parameter ω, which is 2 in the case of the scalar free field theory, is not known in general and has to be determined empirically.

We have encountered the most straightforward methods for simulating the scalar field theory. Most of these methods can easily be generalised to more complicated field theories. Before discussing different methods, we shall analyse the behaviour of the methods presented so far.

15.4.2 The MC algorithms: implementation and results

The implementation of the algorithms presented in the previous sections is straightforward. The reader is encouraged to try coding a few and to check the results given below.

To obtain the renormalised mass and coupling constant, Eqs. (15.53) and (15.54) can be used – however it is nice to measure the full two-point correlation function. This can be found by sampling this function for pairs of points which lie in the same column or in the same row. To obtain better statistics, nonhorizontal and nonvertical pairs can be taken into account as well. To this end we construct a histogram, corresponding to equidistant intervals of the pair separation. We keep two arrays in the program, one for the value of the correlation function, and the other for the average distance r corresponding to each histogram column. At regular time intervals we perform a loop over all pairs of lattice sites. For each pair we calculate the closest distance within the periodic boundary conditions according to the minimum image convention. Suppose this distance is r_{ij}. We calculate to which column this value corresponds, and add the product of the field values at the two sites $\phi_i\phi_j$ to the correlation function array. Furthermore we add r_{ij} to the average distance array. After completing the loop over the pairs, we divide the values in the correlation function array and in the average distance array by the number of pairs that contributed to these values. The final histogram must contain the time averages of the correlation function values thus evaluated – this should be written to a file.

We can now check whether the scalar ϕ^4 theory is renormalisable. This means that if we discretise the continuum field theory using finer and finer grids, the resulting physics should remain unchanged. Equation (15.43) tells us how we should change the various parameters of the theory when changing the grid constant. We now present results for a field theory which on an 8×8 lattice is fixed by the parameter values $m = 0.2$ and $g = 0.04$. Note that both m and g should be close to the critical line (which passes through $m = 0, g = 0$) to obtain long correlation lengths justifying the discretisation. According to (15.43) we use $m = 0.1$ and $g = 0.01$ on a 16×16 lattice, etc. The results are obtained using a heat-bath algorithm using 30 000 steps (8×8) to 100 000 steps (24×24). Figure 15.1 shows the correlation functions for various lattice sizes, obtained using the heat-bath algorithm. The horizontal axis is scaled proportional to the lattice constant (which is obviously twice as large for an 8×8 lattice as for a 16×16 lattice). The vertical axis is scaled for each lattice size in order to obtain the best collapse of the various curves. It is seen that for length scales beyond the lattice constant, scaling is satisfied very well. Only on very small length scales do differences show up, as is to be expected.

The correlation functions obtained from the simulations can be compared with the analytic form, which can be obtained by Fourier transforming

$$\mathcal{G}_{k,-k} = \frac{Z}{4\sum_\mu \sin\frac{k_\mu}{2} + m_{\mathrm{R}}^2} \tag{15.59}$$

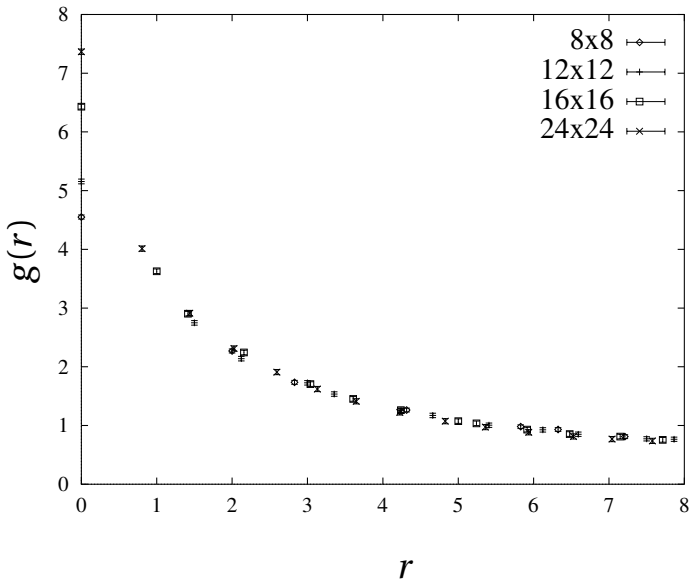

Figure 15.1: The correlation function of the interacting scalar field theory for various lattice sizes. The mass and coupling parameters for the different lattice sizes have been scaled such as to keep the physical lattice size constant. The x-axis has been scaled accordingly. The values have been determined using the histogram method described in the text.

[see Section 15.3 and Eq. (15.53)]. The parameter m_{R} which gives the best match to the correlation function obtained in the simulation (with an optimal value of the parameter Z), is then the renormalised mass. For each of the correlation functions represented in Figure 15.1, parameters Z and m_{R} can be found such that the analytic form lies within the (rather small) error bars of the curves obtained from the simulation. In table 15.1, the values of the renormalised mass as determined using this procedure are compared with those obtained using (15.53). Excellent agreement is found. It is seen that for the larger lattices, the renormalised mass is more or less inversely proportional to the linear lattice size. The physical mass however should be independent of the lattice size. This is because masses are expressed in units of the inverse lattice constant, and the lattice constant is obviously inversely proportional to the linear lattice size L if the lattice represents the same physical volume for different sizes.

The determination of the renormalised coupling constant is difficult. We use Eq. (15.54), but this is subject to large statistical errors. The reason is that the result is the difference of two quantities which are nearly equal, and this difference

Table 15.1: Values of the renormalised mass, obtained from (a) matching the measured correlation function to the analytic form (15.59) and from (b) formula (15.53), for different grid sizes.

L	m	g	$m_{\mathrm{R}}^{(a)}$	$m_{\mathrm{R}}^{(b)}$
8	0.2	0.04	0.374(5)	0.363(4)
12	0.1333	0.01778	0.265(5)	0.265(7)
16	0.1	0.01	0.205(7)	0.204(8)
24	0.06667	0.004444	0.138(4)	0.138(4)

Table 15.2: Values of the renormalised mass and coupling constant obtained from (15.53) and (15.54), for different lattice sizes. Various methods (see later sections) are used.

L	m	g	m_{R}	g_{R}
8	0.05	0.1	0.456(3)	0.20(4)
12	0.03333	0.04444	0.332(3)	0.18(7)
16	0.025	0.025	0.260(3)	0.13(5)
24	0.016667	0.01111	0.184(2)	0.12(6)
32	0.0125	0.00625	0.1466(7)	0.10(4)

is subject to the (absolute) error of these two quantities – hence the *relative* error of the difference becomes very large. The renormalised coupling constant should not depend on the lattice size for large sizes, as it is dimensionless. Table 15.2 shows the results. The errors are rather large and it is difficult to check whether the renormalised coupling constant remains the same indeed, although the data are compatible with a coupling constant settling at a size-independent value of $g \approx 0.11$ for large lattices.

15.4.3 Molecular dynamics

How can we use molecular dynamics for a field theory formulated on a lattice, which has no intrinsic dynamics?[†] The point is that we assign a *fictitious* momentum degree of freedom to the field at each site (the Car-Parrinello method is based on a similar trick – see Chapter 9). As we have seen in Chapter 7 and 8,

[†]The dynamics is here defined in terms of the evolution of the field configuration and not in terms of the time axis of the lattice.

for a dynamical system the probability distribution of the coordinate part can be obtained by integrating out the momentum degrees of freedom, and this should be the desired distribution $e^{-S[\phi]}$. Therefore, we simply add a kinetic energy to the action in order to obtain a classical Hamiltonian (which should not be confused with the field theory's quantum Hamiltonian):

$$\mathscr{H}_{\text{class}} = \sum_n \frac{p_n^2}{2} + S[\phi]. \tag{15.60}$$

Integrating out the momentum degrees of freedom of the classical partition function, we obtain the Boltzmann factor of the action back again (up to a constant):

$$\int [\mathscr{D} p_n]\, e^{-\mathscr{H}_{\text{class}}[p_n,\phi_n]} = \text{Const} \cdot e^{-S[\phi]}. \tag{15.61}$$

15.4.3.1 The Andersen method

The classical Hamiltonian gives rise to classical equations of motion which can be solved numerically. These equations yield trajectories with constant energy (up to numerical errors). We want, however, trajectories representing the canonical ensemble, and in Chapter 8 we studied various methods for obtaining these. In the Andersen refreshed molecular dynamics method, the momenta are refreshed every now and then by replacing them with a new value drawn from a random generator with a Maxwell distribution. In field theories, one often replaces all momenta at the same time with regular intervals between these updates (the method is usually denoted as the *hybrid method*). That is, first the equations of motion are solved for a number of time steps, and then *all* momenta are replaced by new values from the Maxwell random generator. Then the equations of motion are solved again for a number of steps and so on.[13–16] The exact dynamical trajectory plus the momentum update can be considered as one step in a Markov chain whose invariant distribution is the canonical one. We do not obtain the exact dynamical trajectory, but a numerical approximation to it, and the errors made can be corrected for in a procedure which will be discussed in the next section. In Chapter 8 we mentioned that the Andersen method leads indeed to the canonical distribution of the coordinate part. We shall prove this statement now.

First, it is useful to consider 'symmetric' Markov steps: these consist of an integration of the equations of motion over a time $\Delta t/2$, then a momentum refreshing, and then again an integration over a time $\Delta t/2$. Such a step can schematically be represented as follows:

$$\Phi_{\text{i}}, P_{\text{i}} \xrightarrow{\Delta t/2} \Phi_{\text{m}}, P_{\text{m}} \quad \text{Refresh:} \quad \Phi_{\text{m}}, P_{\text{rm}} \xrightarrow{\Delta t/2} \Phi_{\text{f}}, P_{\text{f}}.$$

Energy conservation during the microcanonical trajectories implies

$$H(\Phi_\mathrm{i}, P_\mathrm{i}) = H(\Phi_\mathrm{m}, P_\mathrm{m}); \tag{15.62a}$$

$$H(\Phi_\mathrm{m}, P_\mathrm{rm}) = H(\Phi_\mathrm{f}, P_\mathrm{f}). \tag{15.62b}$$

The steps occur with a probability

$$T(\Phi_\mathrm{i}, P_\mathrm{i} \to \Phi_\mathrm{f}, P_\mathrm{f}) = \delta(\Phi_\mathrm{f} - \Phi_\mathrm{microcanonical}) \exp\left(-P_\mathrm{rm}^2/2\right), \tag{15.63}$$

where the delta-function indicates that Φ_f is uniquely determined by the microcanonical trajectory, which depends of course on the initial configuration $\Phi_\mathrm{i}, P_\mathrm{i}$, the refreshed momentum P_rm, and on the integration time (which is fixed).

The trial steps are ergodic, and the master equation of the Markov chain

$$\sum_{\Phi', P'} \rho(\Phi, P) T(\Phi, P \to \Phi', P') = \sum_{\Phi', P'} \rho(\Phi', P') T(\Phi', P' \to \Phi, P) \tag{15.64}$$

will have an invariant solution. However, the detailed balance condition for this chain is slightly modified. The reason is that we need to use the time-reversibility of the microcanonical trajectories, but this reversibility can only be used when we reverse the momenta. Therefore we have

$$\frac{\rho(\Phi', P')}{\rho(\Phi, P)} = \frac{T(\Phi, P \to \Phi', P')}{T(\Phi', -P' \to \Phi, -P)} \tag{15.65}$$

[note that $\rho(\Phi, P) = \exp\left[-P^2/2 - S(\Phi)\right]$ is symmetric with respect to $P \leftrightarrow -P$]. The transition probability in the denominator of the right hand side corresponds to the step in the numerator traversed backward in time (see the above diagram of a symmetric trial step). The fraction on the right hand side is clearly equal to $\exp\left[(P_\mathrm{mr}^2 - P_\mathrm{m}^2)/2\right]$. Using Eqs. (15.62), it then follows that the invariant distribution is given as $\rho(\Phi, P) = \exp\left[-P^2/2 - S(\Phi)\right]$.

The procedure can be implemented straightforwardly. The equations of motion in the leap-frog form read

$$p_n(t + h/2) = p_n(t - h/2) + hF_n(t); \tag{15.66a}$$

$$\phi_n(t + h) = \phi_n(t) + hp_n(t + h/2), \tag{15.66b}$$

where the force $F_n(t)$ is given by

$$F_n(t) = \sum_\mu \left[\phi_{n+\mu}(t)\right] - (2d + m^2)\phi_n(t) - 2g\phi_n^3(t), \tag{15.67}$$

where $\sum_\mu$ denotes a sum over *all* neighbours. Refreshing the momenta should be carried out with some care. We refresh the momenta at the time steps t for which

the field values ϕ_n are evaluated in the leap-frog algorithm. However, we need the momenta in the leap-frog algorithm at times precisely halfway between these times. Therefore, after the momentum update, we must propagate the momenta over half a time step h:

$$p_n(t + h/2) = p_n(t) + hF_n(t)/2, \tag{15.68}$$

and then the integration can proceed again.

This method contains a tunable parameter: the refresh rate. It turns out[16] that the efficiency has a broad optimum as a function of the refresh rate. Having around fifty steps between the all-momenta updates with a time step $h = 0.1$ is quite efficient. If we refresh after every time step, the system will essentially carry out a random walk in phase space as the small steps made between two refreshings are nearly linear, and the direction taken after each refreshment step is approximately random. If we let the system follow its microcanonical trajectory for a longer time, it will first go to a state which is relatively uncorrelated with respect to the previous one. The momentum refreshings then ensure that the canonical distribution is satisfied; however, the fact that the energy is not conserved, but may change by an amount (on average) of $\mathcal{O}(h^2)$ during the MD trajectory, causes deviations from the canonical distribution of the same order of magnitude.

This method is obviously more efficient than refreshing after each step, as the distance covered by a random walker increases only as the square root of the number of steps made. If we wait too long between two refreshings, the simulation samples only a few different energy surfaces which is not representative for the canonical ensemble. The optimum refresh rate is therefore approximately equal to the correlation time of the microcanonical system.

15.4.3.2 The Metropolis-improved MD method

The leap-frog algorithm introduces systematic errors into the numerical simulation, and the distribution will therefore not sample to the exact one. That is not necessarily a bad thing: we can always write the distribution which is sampled by the MD trajectory as $\exp(-S_{\mathrm{D}}[\phi])$, where the action S_{D} differs by some power of h from the continuum action:[17]

$$S_{\mathrm{D}}[\phi] = S[\phi] + \mathcal{O}(h^k) \tag{15.69}$$

for some positive k. The discrete action may renormalise to a continuum limit with slightly different parameters, but as the behaviour of the model is calibrated in the end by matching calculated physical quantities to the experimental values, our model with discrete time step might still describe the correct continuum limit. Indeed, Batrouni *et al.*[17] show that the discrete time action in the Langevin limit

(i.e. the case in which the momenta are refreshed at every time step, see below) is a viable one at least to first order in h. A problem is that the difference between the discrete and the continuum actions makes it difficult to compare the results of the MD simulation with an MC simulation of the same system with the same values of the parameters.

The discretisation error can be corrected for in exactly the same way as is done in the variational and diffusion quantum Monte Carlo method, see for example the discussion near the end of Section 12.2.5. The idea is to consider the leap-frog MD trajectories as a trial step in a Monte Carlo simulation. The energy before and after this trial step is calculated, and the trial step is accepted with probability $\exp(-\mathcal{H}^{\text{new}}_{\text{class}}+\mathcal{H}^{\text{old}}_{\text{class}})$ (note that $\mathcal{H}_{\text{class}}$ is a classical 'energy' which includes kinetic and potential energy). If it is rejected, the momenta are refreshed once more and the MD sequence starts again. This method combines the Andersen refreshment steps with microcanonical trajectory acceptance/rejection steps. In the previous subsection we saw that the refreshment step satisfies a modified detailed balance condition which ensures the correct (canonical) distribution. Now we show that the microcanonical trajectories plus the acceptance/rejection step, satisfy a similar detailed balance with a canonical invariant distribution.

We write the transition probability in the form of a trial step probability $\omega_{\Phi,P;\Phi',P'}$ and a Metropolis acceptance/rejection probability $A_{\Phi,P;\Phi',P'}$. The trial step probability is determined by the numerical leap-frog trajectory and hence is nonzero only for initial and final values compatible with the leap-frog trajectory. Time-reversibility of the leap-frog algorithm implies that

$$\omega_{\Phi,P;\Phi',P'} = \omega_{\Phi',-P';\Phi,-P}. \tag{15.70}$$

The acceptance probability is given as usual by

$$A_{\Phi,P;\Phi',P'} = \min\left\{1, \exp\left[\mathcal{H}_{\text{class}}(\Phi,P) - \mathcal{H}_{\text{class}}(\Phi',P')\right]\right\}. \tag{15.71}$$

The acceptance step is invariant under $P \leftrightarrow -P$ as the momenta occur only with an even power in the Hamiltonian.

From this, it follows immediately that the modified detailed balance condition holds:

$$\frac{\rho(\Phi',P')}{\rho(\Phi,P)} = \frac{\omega_{\Phi,P;\Phi',P'}A_{\Phi,P;\Phi',P'}}{\omega_{\Phi',-P';\Phi,-P}A_{\Phi',P';\Phi,P}} = \exp\left[\mathcal{H}_{\text{class}}(\Phi,P) - \mathcal{H}_{\text{class}}(\Phi',P')\right]. \tag{15.72}$$

We see that without momentum refreshings, the canonical distribution is a stationary distribution of the Markov process. However, for small time steps in the leap-frog algorithm, the changes in the classical Hamiltonian are very small, and convergence will be extremely slow. That is the reason why these steps are

combined with momentum refreshings, which are compatible with a canonical invariant distribution too, but which cause more drastic changes in the energy. This method is usually called *hybrid Monte Carlo method*.[18]

The important advantage of this Metropolis-improved MD method is that the time step of the leap-frog algorithm can be stretched considerably before the acceptance rate of the Metropolis step drops too low. This causes the correlation time for the 'microcanonical' part, measured in time steps, to be reduced considerably. We have put the quotes around 'microcanonical' because the energy is not conserved very well with a large time step. If the time step is taken too large, the Verlet method becomes unstable (see Section A.7.1.3). In practice one often chooses the time step such that the acceptance rate becomes about 80%, which is on the safe side, but still not too far from this instability limit.

It should be noted that the acceptance rate depends on the difference in the total energy of the system before and after the trial step. The total energy is an extensive quantity: it scales linearly with the volume. This implies that discrete time step errors will increase with volume. To see how strong this increase is,[19] we note that the error in coordinates and momenta after many steps in the leap-frog/Verlet algorithm is of order h^2 per degree of freedom (see problem A.3). This is then the deviation in the energy over the microcanonical trajectory – we shall denote this deviation ΔH_{MD}. The energy differences obtained including the acceptance/rejection step are called ΔH_{MC}, that is, if the trajectory is accepted, ΔH_{MC} is equal to ΔH_{MD}, but if the step is rejected, $\Delta H_{\mathrm{MC}} = 0$. If ΔH_{MD} averaged over all possible initial configurations would vanish, the acceptance rate would always be larger than 0.5, as we would have as many positive as negative energy differences (assuming that the positive differences are on average not much smaller or larger than the negative ones), and all steps with negative and some of the steps with positive energy difference would be accepted. However, the net effect of the acceptance/rejection step is to *lower* the energy, and since the energies measured with this step included remain on average stationary, $\langle \Delta H_{\mathrm{MD}} \rangle$ must be positive. The fact that the energy remains stationary implies that $\langle \Delta H_{\mathrm{MC}} \rangle = 0$ and this leads to an equation for $\langle \Delta H_{\mathrm{MD}} \rangle$:

$$\langle \Delta H_{\mathrm{MC}} \rangle = 0 = \frac{\sum_{\{\Phi,P\}} P_{\mathrm{acc}}\,(\Delta H_{\mathrm{MD}})\,\Delta H_{\mathrm{MD}}}{\sum_{\{\Phi,P\}}}. \tag{15.73}$$

Using $P_{\mathrm{acc}} = \min[1, \exp(-\Delta H_{\mathrm{MD}})]$, and expanding the exponent, we find

$$0 = \langle \Delta H_{\mathrm{MD}} \rangle - \left\langle \theta(\Delta H_{\mathrm{MD}})(\Delta H_{\mathrm{MD}})^2 \right\rangle \tag{15.74}$$

where the theta-function restricts ΔH_{MD} to be positive: $\theta(x) = 0$ for $x < 0$ and 1 for $x > 0$. We see that $\langle \Delta H_{\mathrm{MD}} \rangle$ is indeed positive and we furthermore conclude

that $\langle \Delta H_{\mathrm{MD}} \rangle = \mathcal{O}(h^4 V)$ for of the order of V degrees of freedom. For the average acceptance value we then find

$$\langle P_{\mathrm{acc}} \rangle = \left\langle \min\left(1, e^{-\Delta H_{\mathrm{MD}}}\right) \right\rangle \approx e^{-\langle \Delta H_{\mathrm{MD}} \rangle} = e^{-\alpha h^4 V} \tag{15.75}$$

where α is of order one. Therefore, in order to keep the acceptance rate constant when increasing the volume, we must decrease h according to $V^{-1/4}$, which implies a very favourable scaling.

15.4.3.3 The Langevin method

Refreshing the momenta after every MD step leads to a Langevin-type algorithm. Langevin algorithms have been discussed in Section 8.8 and in Section 12.2.4. In Section 8.8 we applied a Gaussian random force at each time step. In the present case we assign Gaussian random values to the momenta at each time step as in Section 12.2.4. In that case the two steps of the leap-frog algorithm can be merged into one, leading to the algorithm:

$$\phi_n(t+h) = \phi_n(t) + \frac{h^2}{2} F_n(t) + h R_n(t). \tag{15.76}$$

The random numbers R_n are drawn from a Gaussian distribution with a width of 1 – it is a Gaussian momentum, not a force (hence the pre-factor h instead of h^2). Comparing the present approach with the Fokker-Planck equation discussed in Section 12.2.4, we see that when we take ρ of the Fokker-Planck equation (12.42) equal to $\exp(-S[\phi])$, Eq. (12.46) reduces to (15.76) if we put $\Delta t = h^2$. This then shows immediately that the Langevin algorithm guarantees sampling of the configurations weighted according to the Boltzmann distribution.

An advantage of this algorithm is the memory saving resulting from the momenta not being required in this algorithm but, as explained in the previous section, the method is not very efficient because the system performs a random walk through phase space. The reason why we treat this method as a separate one here is that there exists an improved version of it which is quite efficient.[17] We shall discuss this algorithm in Section 15.5.5.

15.4.3.4 Implementation

All the MD algorithms described can be implemented without difficulty. The details of the leap-frog and Langevin algorithm can be found in Chapter 8. Moreover, calculation of the correlation function is described in Section 15.4.2. The programs can all be tested using the results presented in that section.

15.5 Reducing critical slowing down

As we have already seen in Section 7.3.2, systems close to the critical point suffer from *critical slowing down*: this is the phenomenon that the correlation time τ diverges as a power of the correlation length. This renders the calculation of the critical properties very difficult, which is quite unfortunate as these properties are usually of great interest: we have seen in this chapter that lattice field theories must be close to a critical point in order to give a good description of the continuum theory. In statistical mechanics, critical properties are studied very often to identify the critical exponents for various universality classes.

For most systems and methods, the critical exponent z, defined by

$$\tau = \xi^z, \tag{15.77}$$

is close to 2. For Gaussian models, the value $z = 2$ of the critical exponent is related to the convergence time of the simple Poisson solvers, which can indeed be shown to be equal to 2 (see Section A.7.2.2). The value of 2 is related to the fact that the vast majority of algorithms used for simulating field theories are *local*, in the sense that only a small number (mostly one) of degrees of freedom is changed at a step. For systems characterised by domain walls (e.g. the Ising model), the exponent 2 can be guessed by a crude heuristic argument. The major changes in the system configuration take place at the domain walls, as it takes less energy to move a wall than to create new domains. In one sweep, the sites neighbouring a domain wall have on average been selected once. The domain wall will therefore move over a distance 1. But its motion has a random walk nature. To change the configuration substantially, the domain wall must move over a distance ξ, and for a random walk this will take of the order of ξ^2 steps.

Over the last ten years or so, several methods have been developed for reducing the correlation time exponent z. Some of these methods are tailored for specific classes of models, such as the Ising and other discrete spin models. All methods are variations of either the Metropolis method, or of one of the MD methods discussed in the previous section. In this section we shall analyse the different methods in some detail. Some methods are more relevant to statistical mechanics, such as those which are suitable exclusively for Potts models, of which the Ising model is a special case, but we treat them in this chapter because the ideas behind the methods developed for field theories and statistical mechanics are very similar.

As the local character of the standard algorithms seems to be responsible for the critical slowing down present in the standard methods, the idea common to the methods to be discussed is to update the stochastic variables in a *global* fashion, that is, all in one step. How this is done can vary strongly from one method to the other, but the underlying principle is the same for all of them.

15.5.1 The Swendsen-Wang method

We start with the cluster method of Swendsen and Wang (SW)[20] and explain their method for the Ising model in d dimensions, discussed already in Section 7.2.2 and 10.3.1. The SW method is a Monte Carlo method in which the links, rather than the sites, of the Ising lattice are scanned in lexicographic order. For each link there are two possibilities:

1. The two spins connected by this link are opposite. In that case the interaction between these spins is deleted.

2. The two spins connected by the link are equal. In that case we either delete the bond or 'freeze' it, which means that the interaction is made infinitely strong. Deletion occurs with probability $p_{\mathrm{d}} = e^{-2\beta J}$ and freezing with probability $p_{\mathrm{f}} = 1 - p_{\mathrm{d}}$.

This process continues until we have visited every link. In the end we are left with a model in which all bonds are either deleted of 'frozen', that is, their interaction strength is either 0 or ∞. This means that the lattice is split up in a set of disjoint clusters and within each cluster the spins are all equal. This model is simulated trivially by assigning at random a new spin value + or − to each cluster. Then the original Ising bonds are restored and the process starts again, and so on.

Of course we must show that the method does indeed satisfy the detailed balance condition. Before doing so, we note that the method leads indeed to a reduction of the dynamic critical exponent z of the two-dimensional Ising model to the value 0.35 presented by SW,[†] which is obviously an important improvement with respect to the value $z = 2.125$ for the standard MC algorithm. The reason why the method works is that flipping blocks involves flipping many spins in one step. In fact, the Ising (or more generally, the Potts model) can be mapped on a cluster model, where the distribution of clusters is the same as for the SW clusters.[22] The average linear cluster size is proportional to the correlation length, and this will diverge at the phase transition. Therefore, the closer we are to the critical point, the larger the clusters are and the efficiency will increase accordingly.

In the Swendsen-Wang method, any configuration can be reached from any other configuration, because there is a finite probability that the lattice is partitioned into L^d single-spin clusters which are then given values + and − at random. Furthermore it is clear that the method does not generate periodicities in time and it remains to be shown that the SW method satisfies detailed balance. We do this by induction. We show that the freezing/deleting process for some arbitrary bond does not destroy

[†]From a careful analysis, Wolff[21] has found exponents $z = 0.2$ and $z = 0.27$ for the 2D Ising model, depending on the physical quantity considered.

detailed balance, so carrying out this process for every bond in succession does not do so either.

Every time we delete or freeze a particular bond ij we change the Hamiltonian of the system:

$$H \to H_0 + V_{ij}, \tag{15.78}$$

H is the Hamiltonian in which the bond is purely Ising-like. H_0 is the Hamiltonian without the interaction of the bond ij and V_{ij} represents an interaction between the spins at i and j which is either ∞ (in the case of freezing) or 0 (if the bond has been deleted) – the remaining bonds do not change. We write the detailed balance condition for two arbitrary configurations S and S' for the system with Hamiltonian H as follows:

$$\frac{T(S \to S')}{T(S' \to S)} = e^{-\beta[H(S')-H(S)]} = \frac{T_0(S \to S')}{T_0(S' \to S)} e^{-\beta J(s_i' s_j' - s_i s_j)}, \tag{15.79}$$

where T_0 is the transition probability for the Hamiltonian H_0 and we have explicitly split off the contribution from the bond ij. In the last equality we have used the detailed balance condition for the system with Hamiltonian H_0.

In the SW algorithm, we must decide for a bond ij whether we delete or freeze this bond. The transition probability of the system after this step can be written as

$$T(S \to S') = T_\mathrm{f}(S \to S') P_\mathrm{f}(S) + T_\mathrm{d}(S \to S') P_\mathrm{d}(S). \tag{15.80}$$

Here, $P_{\mathrm{d,f}}(S)$ is the probability that we delete (d) or freeze (f) the bond ij in spin configuration S. $T_\mathrm{d}(S \to S')$ is the transition probability with a deleted bond, and therefore $T_\mathrm{d} = T_0$, and T_f is the transition probability when the bond is frozen. The latter is equal to T_0 in the case that the spins s_i, s_j are equal in both S and S' and it is zero in the case that they are unequal in S' (they must be equal in S, otherwise they could not have been frozen).

Let us consider the detailed balance condition for the transition probability in (15.80):

$$\frac{T(S \to S')}{T(S' \to S)} = \frac{T_\mathrm{f}(S \to S') P_\mathrm{f}(S) + T_\mathrm{d}(S \to S') P_\mathrm{d}(S)}{T_\mathrm{f}(S' \to S) P_\mathrm{f}(S') + T_\mathrm{d}(S' \to S) P_\mathrm{d}(S')} = e^{-\beta[H(S')-H(S)]}. \tag{15.81}$$

We show that this condition is indeed satisfied, using (15.79). Let us assume that s_i and s_j are equal in both S and S'. In that case $P_\mathrm{f}(S) = 1 - \exp(-2\beta J)$ and $P_\mathrm{d} = \exp(-2\beta J)$ respectively and we have

$$\frac{T(S \to S')}{T(S' \to S)} = \frac{T_\mathrm{f}(S \to S') \left[1 - \exp(-2\beta J)\right] + T_0(S \to S') \exp(-2\beta J)}{T_\mathrm{f}(S' \to S) \left[1 - \exp(-2\beta J)\right] + T_0(S' \to S) \exp(-2\beta J)}. \tag{15.82}$$

Since the pair $s_i s_j$ is equal in both S and S', the transition probability $T_\mathrm{f} = T_0$ and we see that (15.79) holds indeed for the transition probability after the SW step.

If before and after the step the spins s_i and s_j are unequal, T_f vanishes in both numerator and denominator, and it is clear that (15.79) holds in this case too. Suppose $s_i = s_j$ in S and that the corresponding pair s'_i, s'_j in S' is unequal. In that case we have

$$\frac{T(S \to S')}{T(S' \to S)} = \frac{T_\mathrm{f}(S \to S')\left[1 - \exp(-2\beta J)\right] + T_0(S \to S')\exp(-2\beta J)}{T_0(S' \to S)}. \qquad (15.83)$$

The denominator in the right hand side contains only the term with a deleted bond because starting from the configuration S' in which s'_i and s'_j are unequal, we can only delete the bond. The transition probability $T_\mathrm{f}(S \to S')$ occurring in the numerator obviously vanishes, and we see that also for this case detailed balance, Eq. (15.79), is again satisfied.

It is instructive to code the Swendsen-Wang method. First a sweep through the lattice is performed in which all the bonds are either frozen or deleted. This poses no difficulties. Then the clusters must be identified. This can be done using 'back-tracking' and is most conveniently coded recursively. It works as follows. A routine BackTrack(x, y) is written, which scans the cluster containing the site given by the Cartesian (integer) components (x, y). Start at site (x, y) and check whether this site has already been visited. If this is not the case, leave a flag there as a mark that the cluster site has now been visted, and scan the neighbouring sites in a similar way by recursive calls. The resulting routine looks more or less as follows (for $d = 2$):

```
ROUTINE BackTrack(x, y)
    IF NOT Visited (x,y) THEN
          Mark (x, y) as being visited;
          IF (Frozen(x, y, x + 1, y)) THEN
              BackTrack(x + 1, y);
          IF (Frozen(x, y, x, y + 1)) THEN
              BackTrack(x, y + 1);
          IF (Frozen(x, y, x − 1, y)) THEN
              BackTrack(x − 1, y);
          IF (Frozen(x, y, x, y − 1)) THEN
              BackTrack(x, y − 1);
    END IF
END BackTrack.
```

Frozen$(x1, y1, x2, y2)$ is a boolean function which returns TRUE if the nearest neighbour bond between $(x1, y1)$ and $(x2, y2)$ is frozen and FALSE otherwise.

Periodic boundaries should be implemented using a modulo operator or function, and it is convenient to decide before scanning the cluster whether it is going to be flipped and, if yes, to do so during the recursive scanning (alongside putting the Visited flag). On exit, the cluster is scanned and all its sites marked as visited. In a sweep through all values i and j, all clusters will be found in this way and it is to be noted that the computer time needed to scan a cluster in the backtrack algorithm scales linearly with the cluster volume (area).

Another algorithm for detecting all the clusters in the system is that of Hoshen and Kopelman. This algorithm does not use recursion. It scales linearly with the lattice size and it is more efficient than back-tracking (30–50%) but it is somewhat more difficult to code. Details can be found in the literature.[23]

The time scaling exponent z can be determined from the simulations. Note that the time correlation of the magnetisation is useless for this purpose as the clusters are set to arbitrary spin values after each sweep, so that the magnetisation correlation time is always of order 1. Therefore, we consider the time correlation function of the (unsubtracted) susceptibility per site. This is defined as

$$\chi = \frac{1}{L^{2d}} \left\langle \left(\sum_i s_i \right)^2 \right\rangle. \tag{15.84}$$

Its time correlation function is

$$C_\chi(k) = \frac{\sum_{n=1}^{N} \chi_{n+k} \chi_n}{\sum_{n=1}^{N} \chi_n^2} \tag{15.85}$$

where the indices n and k are 'time' indices, measured in MC steps per spin.

The susceptibility can be determined directly from the lattice configuration after each step using (15.84), but it is possible to obtain a better estimate by realising that when the system is divided up into clusters c of area N_c, the average value of χ is given by

$$\chi = \frac{1}{L^{2d}} \left\langle \left(\sum_c N_c s_c \right)^2 \right\rangle \tag{15.86}$$

where s_c is the spin value of cluster c. We can write this as

$$\chi = \frac{1}{L^{2d}} \left\langle \sum_c N_c s_c \sum_{c'} N_{c'} s_{c'} \right\rangle, \tag{15.87}$$

and by summing over $s_c = \pm 1$ for all the clusters we obtain the average of this value for all possible cluster-spin configurations. Then only the terms $c = c'$ survive and we are left with

$$\chi = \frac{1}{L^{2d}} \left\langle \sum_c N_c^2 \right\rangle. \tag{15.88}$$

This is the so-called 'improved estimator' for the unsubtracted susceptibility.[24, 25] This estimator gives better results because the average over all possible cluster-spin configurations is built into it.

The correlation time can be determined from the values of χ at the subsequent MC steps in the usual way (see Section 7.4). For a detailed analysis of the dynamic exponent for various cases, see Ref. 21.

– Programming exercise –

Code the SW algorithm for the two-dimensional Ising model. Determine the time relaxation exponent and compare this with the value found for the single-spin flip algorithm.

Wolff[24] has carried out the cluster algorithm in the microcanonical ensemble, using a microcanical MC method proposed by Creutz.[26] He fixed the number of unequal bonds to half the number of total bonds and found considerable improvement in the efficiency.

15.5.2 Wolff's single cluster algorithm

Wolff[27] has proposed a different cluster method for eliminating critical slowing down for Potts spin systems, and an extension of this method and the SW method to a special class of continuous spin models. We start with Wolff's modification of the SW method for the Ising model. In Wolff's method, at each step a *single* cluster is generated as opposed to the SW model, in which the entire lattice is partitioned into clusters. The single cluster is constructed according to the same rules as the SW clusters. We start with a randomly chosen spin and consider its neighbours. Only equal neighbours can be linked to the cluster by freezing the bonds between them – this happens with probability $1 - e^{-2\beta J}$. The cluster is extended in this way until no more spins are added to it. Then all the spins in the cluster are flipped.

It will be clear that the cluster generated in this way is a SW cluster and therefore the method satisfies detailed balance. The difference between the two methods is that the selection of the cluster to be grown can be viewed as throwing a dart at the lattice[21] with equal probability to hit any of the sites – the probability of hitting a SW cluster (SWC) of size N_{SWC} is N_{SWC}/L^d (for d dimensions), thereby favouring large clusters to be grown. Because of this preference for large clusters it is expected that the single cluster version changes the configuration on average more drastically in the same amount of time and that statistically independent configurations are generated in fewer steps. This turns out to be the case in the 3D Ising model, where the single cluster algorithm yields time correlation exponents 0.28 or 0.14 (depending on the correlation function studied) as opposed to 0.5 for

the SW algorithm – for the 2D Ising model only a small increase in efficiency has been measured.[21]

It is convenient to generate the clusters in a recursive way. Each MC step consists of selecting a random site (*Location*) on the lattice – *ClusterSpin* is minus the spin at this location (the spins are flipped when added to the cluster). The algorithm is then as follows:

```
ROUTINE GrowCluster(Location, ClusterSpin):
    Flip Spin at Location;
    Mark Spin as being added to Cluster;
    IF right-hand neighbour not yet added THEN
        TryAdd(RightNeighbour, ClusterSpin);
    ...Similar for other neighbours...
END GrowCluster.

ROUTINE TryAdd (Location, ClusterSpin):
    Determine Spin at Location;
    IF Spin opposite to ClusterSpin THEN
        IF Random number < 1 – e^{-2J} THEN
            GrowCluster(Location, ClusterSpin);
        END IF;
    END IF;
END TryAdd.
```

Measuring correlation times requires some care, as a step in the Wolff algorithm consists of flipping one cluster instead of (on average) half of the total number of spins in the lattice as in the SW algorithm. The correlation time $\bar{\tau}_{\mathrm{W}}$ expressed in numbers of single cluster flips must therefore be translated into the single cluster correlation time τ_{W} expressed in SW time steps:

$$\tau_{\mathrm{W}} = \bar{\tau}_{\mathrm{W}} \frac{\langle N_{1\mathrm{C}} \rangle}{L^d}. \tag{15.89}$$

The average single cluster size $\langle N_{1\mathrm{C}} \rangle$ occurring in the right hand side is the improved estimator for the (unsubtracted) susceptibility per site:

$$\langle N_{1\mathrm{C}} \rangle = \left\langle \frac{N_{\mathrm{SWC}}}{L^d} N_{\mathrm{SWC}} \right\rangle = \chi. \tag{15.90}$$

This formula can be understood by realising that the probability of generating a SW cluster of size N_{SWC} in the single cluster algorithm is equal to N_{SWC}/L^d. To evaluate the average cluster size we must multiply this probability with N_{SWC} and take the expectation value of the result.

– Programming exercise –

Implement Wolff's single cluster algorithm and compare the results with the SW algorithm – see also Ref. 21.

In many statistical spin systems and lattice field theories, the spins are not discrete but they assume continuous values. Wolff's algorithm was formulated for a particular class of such models, the O(N) models. These models consist of spins, which are N-dimensional unit vectors, on a lattice. Neighbouring spins $\mathbf{s}_i, \mathbf{s}_j$ interact – the interaction is proportional to the scalar product $\mathbf{s}_i \cdot \mathbf{s}_j$. An example which is relevant to many experimental systems (superfluid and superconducting materials, arrays of coupled Josephson junctions ...) is the O(2), or XY-model, in which the spins are unit vectors $\mathbf{s}_i$ lying in a plane, so that they can be characterised by their angle θ_i with the x-axis, $0 \le \theta_i < 2\pi$.

For simulations, it is important that relevant excitations in $O(N)$ models are smooth variations of the spin orientation over the lattice (except near isolated points – see below). This implies that changing the value of a single angle θ_i somewhere in the lattice by an amount of order 1 is likely to lead to an improbable configuration – hence the acceptance rate for such a trial change is on average very small. The only way of achieving reasonable acceptance rates for changing a single spin is by restricting the variation in the orientation of the spin allowed in a trial step considerably. This however will reduce the efficiency because many MC steps are then needed to arrive at statistically independent configurations. A straightforward generalisation of the SW or single cluster algorithm in which all spins in some cluster are reversed is bound to fail for the same reason, as this destroys the smoothness of the variation of the spins at the cluster boundary.

Wolff[27] has proposed a method in which the spins in a cluster are modified to an extent depending on their orientation. It turns out that his method can be formulated as an embedding of an Ising model into an $O(N)$ model.[28] First a random unit vector $\mathbf{u}$ is chosen. Every spin $\mathbf{s}_i$ is then split into two components: the component along $\mathbf{u}$ and that perpendicular to $\mathbf{u}$:

$$\mathbf{s}_i^{\parallel} = (\mathbf{s}_i \cdot \mathbf{u})\mathbf{u} \tag{15.91a}$$

$$\mathbf{s}_i^{\perp} = \mathbf{s}_i - \mathbf{s}_i^{\parallel}. \tag{15.91b}$$

We keep $\mathbf{s}_i^{\perp}$ and $|\mathbf{s}_i^{\parallel}|$ fixed – the only freedom left for the $O(N)$ spins is to flip their parallel component:

$$\mathbf{s}_i = \mathbf{s}_i^{\perp} + \epsilon_i \, |\mathbf{s}_i^{\parallel}| \mathbf{u}, \qquad \epsilon_i = \pm 1. \tag{15.92}$$

A flip in the sign ϵ_i corresponds to a reflection with respect to the hyperplane perpendicular to $\mathbf{u}$ (see Figure 15.2). The interaction of the model with the

restriction on the fluctuations that only flips of the parallel components are allowed, can now be described entirely in terms of the ϵ_i:

$$\mathscr{H}[\epsilon_i] = \sum_{\langle ij \rangle} J_{ij}\, \epsilon_i \epsilon_j \tag{15.93a}$$

$$J_{ij} = J|\mathbf{s}_i^{\parallel}|\,|\mathbf{s}_j^{\parallel}|. \tag{15.93b}$$

This Ising Hamiltonian is now simulated using the single cluster or the SW algorithm. After choosing the unit vector $\mathbf{u}$, we calculate the ϵ_i for the actual orientations of the spins and then we allow for reflections of the $\mathbf{s}_i$ (that is, for spin flips in the ϵ_i system).

This method is more efficient than the standard single-spin update method because large clusters of spins are flipped at the same time. But why is the acceptance rate for such a cluster update not exceedingly small? The point is that the amount by which a spin changes, depends on its orientation (see Figure 15.2): for a spin more or less perpendicular to $\mathbf{u}$, the change in orientation is small. This translates itself into the coupling J_{ij} being small for spins $\mathbf{s}_i$, $\mathbf{s}_j$ nearly

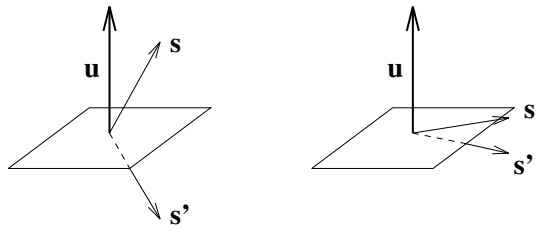

Figure 15.2: Spin flips in the Wolff algorithm for the $O(3)$ model.

perpendicular to $\mathbf{u}$. For spins parallel to $\mathbf{u}$, the coupling constant J_{ij} is large and these spins will almost certainly be frozen to the same cluster. The cluster boundaries will be the curves [in two dimensions, and (hyper)surfaces in higher dimensions] on which the spins $\mathbf{s}_i$ are more or less perpendicular to $\mathbf{u}$. In other words, if we provide a direction $\mathbf{u}$, the algorithm will find an appropriate cluster boundary such that the spin reflection does not require a vast amount of energy. Therefore, the acceptance rate is still appreciable.

The procedure is ergodic as every unit vector $\mathbf{u}$ can in principle be chosen and there is a finite probability that the cluster which is to be swapped consists of a single spin – the isolated spin-update method is therefore included in the new algorithm. Detailed balance is satisfied because the new Ising Hamiltonian (15.93) is exactly equivalent to the original $O(N)$ Hamiltonian under the restriction that only the reflection steps described are allowed in the latter.

The implementation of the method for the two-dimensional XY model proceeds along the same lines as described above for the Ising model. Apart from selecting a random location from which the cluster will be grown, a unit vector $\mathbf{u}$ must be chosen, simply by specifying its angle with the X-axis. Each spin is flipped when added to the cluster. If we try to add a new spin $\mathbf{s}_i$ to the cluster (in routine 'TryAdd'), we need the spin value of its neighbour $\mathbf{s}_j$ in the cluster – the freezing probability P_f is then calculated as

$$P_\mathrm{f} = 1 - \min\left\{1, \exp[\beta \mathbf{s}_i \cdot (\mathbf{s}_j - \mathbf{s}'_j)]\right\} = 1 - \min\left\{1, \exp[2\beta \mathbf{s}_i \cdot \mathbf{s}_j]\right\} \tag{15.94}$$

(note that the cluster spin $\mathbf{s}_i$ has already been flipped, in contrast to $\mathbf{s}_j$). The spin $\mathbf{s}_j$ is then added to the cluster with this freezing probability. Instead of considering continuous angles between 0 and 2π, it is possible to consider an *n-state clock model*, which is an XY model with the restriction that the angles allowed for the spins assume the values $2j\pi/n$, $j = 0, \ldots, n-1$;[29] see also Section 12.6. At normal accuracies, the discretisation of the angles will not be noticed for n greater than about 20. The cosines and sines needed in the program will then assume n different values only and these can be calculated in the beginning of the program and stored in an array.

– Programming exercise –

Write a Monte Carlo simulation program for the XY model, using Wolff's cluster algorithm. If the program works correctly, it should be possible to detect the occurrence of the so-called Kosterlitz-Thouless phase transition (note that this occurs only in two dimensions). This is a transition which has been observed experimentally in helium-4 films[30] and Josephson junction arrays.[31] We shall briefly describe the behaviour of the XY model.

Apart from excitations which are smooth throughout the lattice – *spin-waves* – the XY model exhibits so-called *vortex excitations*. A pair of vortices is shown in Figure 15.3. The vortices can be assigned a *vorticity* which is roughly the 'winding' number of the spins along a closed path around the vortex – the vorticity assumes values $2\pi, -2\pi$ (for higher temperatures also 4π etc. can occur). An isolated vortex requires an amount of energy which scales logarithmically with the lattice size and is hence impossible in a large lattice at finite temperature. However, vortex pairs of opposite vorticity are possible; their energy depends on the distance between the vortices and is equal to the Coulomb interaction (which is proportional to $\ln R$ for two dimensions) for separations R larger than the lattice constant. The system can only contain equal numbers of positive and negative vortices. At low temperatures the vortices occur in bound pairs of opposite vorticity (to be compared

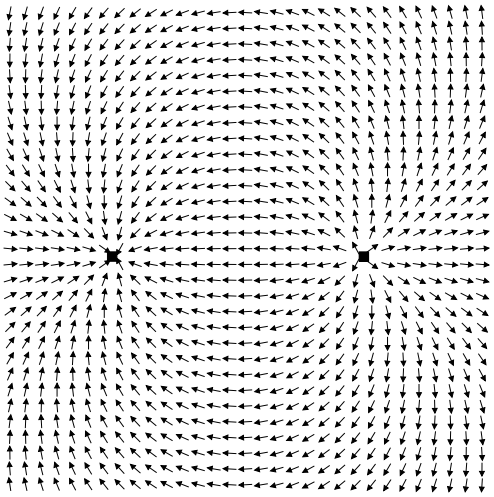

Figure 15.3: A pair of vortices, one with positive and one with negative vorticity.

with electrical dipoles), and the spin-waves dominate the behaviour of the model in this phase. It turns out that the correlations are long-ranged:

$$\langle \theta_i - \theta_j \rangle \sim \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|^{x_{\mathrm{T}}}}, \tag{15.95}$$

for large separation $|\mathbf{r}_i - \mathbf{r}_j|$, with a critical exponent x_{T} which varies with temperature. At the KT transition, the dipole pairs unbind and beyond the transition temperature T_{KT} we have a fluid of freely moving vortices (to be compared with a plasma).

Imagine you have an XY lattice with fixed boundary conditions: the spins have orientation $\theta = 0$ on the left hand side of the lattice and you have a handle which enables you to set the fixed value δ of the spin orientation of the rightmost column of XY-spins. Turning the handle from $\delta = 0$ at low temperatures, you will feel a resistance as if it is attached to a spring. This is due to a nonvanishing amount of free energy which is needed to change the orientation of the spins on the right hand column. This excess free energy scales as

$$\Delta F \sim \Gamma \delta^2 \tag{15.96}$$

for small angles δ. At the KT temperature the force needed to pull the handle drops to zero, as the vortex system has melted, which is noticeable through the proportionality constant Γ dropping to zero.

The quantity Γ is called *spin-wave stiffness*[32] or *helicity modulus*. It can be calculated in a system with periodic boundary conditions using the following formula:[32]

$$\Gamma = \frac{J}{2L^2}\left\{\left\langle \sum_{\langle ij\rangle} \cos(\theta_i - \theta_j)\right\rangle - \frac{J}{k_{\mathrm{B}}T}\left\langle\left[\sum_i \sin(\theta_i - \theta_{i+\hat{e}_x})\right]^2\right\rangle - \frac{J}{k_{\mathrm{B}}T}\left\langle\left[\sum_i \sin(\theta_i - \theta_{i+\hat{e}_y})\right]^2\right\rangle\right\}. \quad (15.97)$$

From the Kosterlitz-Thouless theory[29, 33, 34] it follows that the helicity modulus has a universal value $\Gamma = 2k_{\mathrm{B}}T_{\mathrm{KT}}/\pi$ at the KT transition. The drop to zero is smooth for finite lattices but it becomes steeper and steeper with increasing lattice size. Figure 15.4 shows Γ/J as a function of $k_{\mathrm{B}}T/J$. The line $\Gamma/J = 2(k_{\mathrm{B}}T/J)/\pi$ is also shown and it is seen that the intersection of the helicity modulus curve with this line gives the value from which the helicity modulus drops to zero. You can

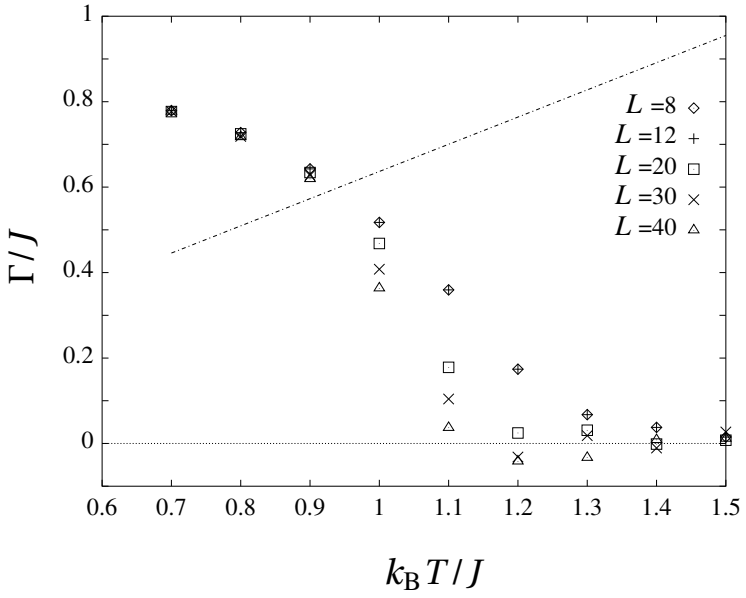

Figure 15.4: The helicity modulus in units of the coupling constant J of the XY model vs. the inverse coupling constant in units of $k_{\mathrm{B}}T$. The intersection of the helicity modulus curves with the straight line gives the value from which the helicity modulus starts dropping to zero.

check your program by reproducing this graph.

Edwards and Sokal[28] have found that for the XY model the dynamic critical exponent in the low-temperature phase is zero or almost zero.

15.5.3 Geometric cluster algorithms

The cluster algorithms described so far flips or rotates spins on a lattice. In fact, Wolff's version of the algorithm for the XY model boils down to flipping an Ising spin. Cluster algorithms strongly rely on a reflection symmetry of the Hamiltonian: flipping all spins does not affect the Hamiltonian. This is the reason why a simple distinction can be made for pairs which may be frozen and those which certainly will not. The same holds for the q-state Potts model: there we have a symmetry under permutation of all spin values. Another way of looking at this is that flipping a large cluster in an Ising model with a magnetic field, yields an energy loss or gain proportional to the volume (surface in two dimensions), which leads to very low acceptance rates in phases with a majority spin. Hence cluster algorithms will be less efficient for such systems.

If we switch on a magnetic field, the spin-flip symmetry is broken and the cluster algorithm can no longer be used. Another problem is that it is not immediately clear whether and how cluster ideas may be generalised to systems which are not formulated on a lattice. A step towards a solution was made by Dress and Krauth,[35] who used *geometric* symmetries to formulate a cluster algorithm for particles moving in a continuum. Usually, a reflection of the particles with respect to a point chosen randomly in the system is used. The interaction between these particles is considered to be a hard-core interaction, but long(er) ranged interaction may also be present. The problem with the algorithm is that the decision to displace a particle is made based on the hard-core part. Other interactions are included in the acceptance criterion, and this leads to many rejections. This problem was solved by Liu and Luijten[36] who take all interactions into account. They start with identifying a random reflection point and then choose an initial particle at random. This and other particles having non-negligible interaction with the first particles, are then candidates to be reflected. This is done one by one, taking all interactions into account, and each time a reflection of a particle would result in a decrease Δ of the energy, the particle is reflected with probability $\exp(-|\Delta|)$ – if the energy increases, the particle is note reflected. This algorithm promises to be valuable for the analysis of dense liquids.

The geometric cluster idea has also been used for spin systems formulated on a lattice.[37] Again, a reflection site is identified at random. Then, for a randomly chosen site i, the spin is exchanged with that of its reflection partner i'. Then each neighbour k of i is investigated. If exchanging the spins at k and k' results in an energy gain Δ, (that is, the total energy decreases), then the move is accepted with

$\exp(-\Delta)$; if this is note the case, k is left unaltered. Then the algorithm proceeds with the neighbours of k just as in Wolff's cluster algorithm. We see that spins are only exchanged in this algorithm, so that the total spin remains constant: the energy change no longer scales with the cluster volume.

15.5.4 The multigrid Monte Carlo method

The multigrid Monte Carlo (MGMC) method,[9, 38, 39] is yet another way of reducing critical slowing down near the critical point. This method is closely related to the multigrid method for solving partial differential equations described in Section A.7.2.5 and readers not familiar with this method should go through that section first – see also problem A.7.

Multigrid ideas can be used to devise a new Monte Carlo algorithm which reduces critical slowing down by moving to coarser and coarser grids and updating these in an MC procedure with a restricted form of the Hamiltonian.

To be specific, let us start from a grid at level l; a field configuration on this grid is called ψ. The Hamiltonian on this grid is called $H_l[\psi]$. The coarse grid is the grid at level $l-1$, and configurations on this coarse grid are denoted by ϕ. Now consider the prolongation operation $P_{l,l-1}$ described in Section A.7.2.5, which maps a configuration ϕ on the coarse grid to a configuration ψ on the fine grid by copying the value of ϕ on the coarse grid to its four nearest neighbours on the fine grid:

$$P_{l,l-1} : \phi \to \psi; \tag{15.98a}$$

$$\psi(2i+\mu, 2j+\nu) = \phi(i,j), \tag{15.98b}$$

where μ and ν are ± 1. We consider now a restricted Hamiltonian $H_{l-1}[\delta\phi]$, which is a function of the coarse grid configuration $\delta\phi$, depending on the fine grid configuration ψ which is kept fixed:

$$H_{l-1}[\delta\phi] = H_l[\psi + P_{l,l-1}(\delta\phi)]. \tag{15.99}$$

We perform a few MC iterations on this restricted Hamiltonian and then we go to the coarser grid at level $l-2$. This process is continued until the lattice consists of a single site, and then we go back by copying the fields on the coarser grid sites to the neighbouring sites of the finer grids, after which we perform again a few MC steps, and so on.

The algorithm reads, in recursive form:

ROUTINE MultiGridMC(l, ψ H_l)
  Perform a few MC sweeps: $\psi \to \psi'$;

IF (l >0) THEN
Calculate the form of the Hamiltonian
on the coarse grid: $H_{l-1}[\delta\phi] = H_l[\psi' + P_{l,l-1}(\delta\phi)]$;
Set $\delta\phi$ equal zero;
MultiGridMC($l-1$, $\delta\phi$, H_{l-1});
ENDIF;
$\psi'' = \psi' + P_{l,l-1}\delta\phi$;
Perform a few MC sweeps: $\psi'' \to \psi'''$;
END MultiGridMC.

The close relation to the multigrid algorithm for solving Poisson's equation, given in Section A.7.2.5 is obvious.

The MC sweeps consist of a few Metropolis or heat bath iterations on the fine grid field ψ. This step is ergodic as the heat bath and Metropolis update is ergodic. Note that the coarse grid update in itself is not ergodic because of the restriction imposed on fine grid changes (equal changes for groups of four spins) – the Metropolis or heat bath updates are essential for this property.

We should also check that the algorithm satisfies detailed balance. Again, the Metropolis or heat bath sweeps respect detailed balance. The detailed balance requirement for the coarse grid update is checked recursively. A full MCMG step satisfies detailed balance if the coarse grid update satisfies detailed balance. But the coarse grid update satisfies detailed balance if the coarser grid update satisfies detailed balance. This argument is repeated until we reach the coarsest level ($l = 1$). But at this level we perform only a few MC sweeps, which certainly satisfy detailed balance. Therefore, the full algorithm satisfies detailed balance.

There is one step which needs to be worked out for each particular field theory: constructing the coarse Hamiltonian H_{l-1} from the fine one, H_l. We do not know *a priori* whether new interactions, not present in the fine Hamiltonian, will be generated when constructing the coarse one. This often turns out to be the case. As an example, consider the scalar interacting ϕ^4 field theory. The terms ϕ^2 and ϕ^4 generate linear and third powers in ϕ when going to the coarser grid. Moreover, the Gaussian coupling $(\phi_n - \phi_{n+\mu})^2$ generates a term $\phi_n - \phi_{n+\mu}$. Therefore, the Hamiltonians which we must consider have the form:

$$\mathscr{H}[\psi] = \frac{1}{2}\left\{ \sum_{\langle nn' \rangle} \left[J_{n,n'}(\psi_n - \psi_{n'})^2 + \sum_{\mu} K_{n,n'}(\psi_n - \psi_{n'}) \right] + \sum_n \left[L_n\psi_n + M_n\psi^2 + T_n\psi_n^3 + G_n\psi_n^4 \right] \right\}. \quad (15.100)$$

Restricting this Hamiltonian to a coarser grid leads to new values for the coupling constants.

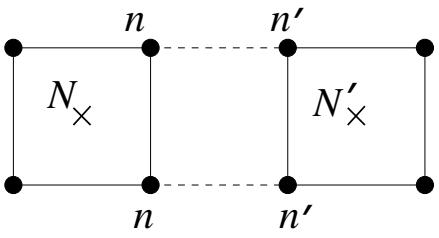

Figure 15.5: Two neighbouring blocks on a fine lattice with coarse lattice sites N and N'.

We denote the sites of the new grid by N, N'. Furthermore, $\sum_{nn'|NN'}$ denotes a sum over sets n, n' of neighbouring points, which belong to different neighbouring blocks of four sites belonging to N and N' respectively as in Figure 15.5. Finally, $\sum_{n|N}$ denotes a sum over the fine grid sites n belonging to the block N. With this notation, the new coupling constants on the coarse grid can be written in terms of those on the fine grid:

$$
\begin{aligned}
J_{NN'} &= \sum_{nn'|NN'} J_{nn'}; \qquad K_{NN'} = \sum_{nn'|NN'} \left[K_{nn'} + 2J_{nn'}\left(\psi_n - \psi_{n'}\right)\right]; \\
L_N &= \sum_{n|N} \left(L_n + 2M_n\psi_n + 3T_n\psi_n^2 + 4G_n\psi_n^3\right); \\
M_N &= \sum_{n|N} \left(M_n + 3T_n\psi_n + 6G_n\psi_n^2\right); \\
T_N &= \sum_{n|N} \left(T_n + 4G_n\psi_n\right); \qquad G_N = \sum_{n|N} G_n.
\end{aligned}
\tag{15.101}
$$

With this transformation, the MCMG method can be implemented straightforwardly. It can be shown that critical slowing down is completely eliminated for Gaussian type actions, so it will work very well for the ϕ^4 theory close to the Gaussian fixed point. However, the ϕ^4 theory has more than one critical point in two dimensions. One of these points has Ising character: for this point, the coefficient of the quadratic term is negative, whereas the coefficient g of ϕ^4 is positive. This means that the field has two opposite minima. For this model, the MCMG method does not perform very well. This can be explained using a heuristic argument. Suppose the field assumes values very close to +1 or −1. Consider a block of four spins on the fine lattice which belong to the same coarse lattice site. Adding a nonzero amount, ϕ_N, to these four spins will only be accepted if they are either all equal to −1, so that an amount of 2 can be added, or if they are all equal to +1 so that we can subtract 2 from each of them. The probability that all spins in a block have equal values becomes smaller and smaller when coarsening the lattice

more and more, so the efficiency of the MCMG method is degraded severely for this case. However, it turns out to be still more efficient by a factor of about 10 with respect to the standard heat bath method.

15.5.5 The Fourier-accelerated Langevin method

We have encountered the Langevin method for field theories in Section 15.4.3.3. This method suffered from slow convergence as a result of small, essentially random, steps being taken, causing the system to perform a random walk in phase space. In 1985, Batrouni *et al.*[17] proposed a more efficient version of the Langevin method in which the fields are updated globally. This is done by Fourier transforming the field, and then applying the Langevin method to the Fourier modes, rather than to the local fields. That this is a valid approach can be seen as follows. We have seen that MD methods can be applied to fields after assigning fictitious momenta to the field variables. In the MD method we have assigned a momentum p_n to each field variable ϕ_n. It is, however, perfectly possible to assign the momenta not to each individual field variable, but to linear combinations of the field variables. After integrating out the momenta we shall again find a Boltzmann distribution of the field variables.

In addition we have the freedom to assign a different time step to each linear combination of field variables. As we have seen in Section 9.3.2, this is equivalent to changing the mass associated with that variable, but we shall take the masses all equal to 1, and vary the time step.

In the Fourier-accelerated Langevin method, we assign momenta p_k to each Fourier component ϕ_k of the field. Furthermore, we choose a time step h_k for each k individually. To be specific, we write the action S in terms of Fourier transformed fields, and construct the following classical Hamiltonian expressed in terms of Fourier modes:

$$\mathcal{H}_{\text{Class}} = \sum_k \left\{ \frac{p_k^2}{2} + S[\phi_k] \right\}. \tag{15.102}$$

By integrating out the momenta it is clear that an MD simulation at constant temperature for this Hamiltonian leads to the correct Boltzmann distribution of the field. In the Langevin leap-frog form, the equation of motion reads

$$\phi_k(t + h_k) = \phi_k(t) - \frac{h_k^2}{2} \frac{\partial S[\phi_k(t)]}{\partial \phi_k} + h_k R_k, \tag{15.103}$$

where R_k is the Fourier transform of a Gaussian random number with a variance of 1 (see below). Fourier transforms are obviously carried out using the fast Fourier transform (see Section A.9).

For a free field model, the dynamical system described by the Hamiltonian (15.102) can be solved trivially, as the Hamiltonian does not contain couplings between the different k's. In that case the action can be written as

$$S[\phi_k] = \frac{1}{2L^{2d}} \phi_k K_{k,-k} \phi_{-k}. \tag{15.104}$$

$K_{k,k'}$ is the free field propagator given in (15.49). The Hamiltonian describes a set of uncoupled harmonic oscillators with periods $T_k = 2\pi / \sqrt{K_{k,-k}}$. The algorithm will be unstable when one of the time steps h_k becomes too large with respect to T_k (see Section A.7.1.3). The most efficient choice for the time steps is therefore

$$h_k = \alpha T_k = \alpha \frac{2\pi}{\sqrt{4 \sum_\mu \sin^2 \frac{k_\mu}{2} + m^2}}, \tag{15.105}$$

where α is some given, small fraction, e.g. $\alpha = 0.2$. If we take all the h_k smaller than the smallest period, then the slower modes would evolve at a much smaller rate than the fast modes. By adopting convention (15.105), the slow modes evolve at exactly the same rate as the fast ones. Therefore, critical slowing down will be completely eliminated for the free field model. For the interacting field with a ϕ^4 term present, the time steps are taken according to (15.105), but with the renormalised mass occuring in the denominator.[17]

A remark is in place here. As the method uses finite time steps, it is not the continuum action which is simulated, but the discrete version which deviates to some order of the time steps from the continuum one. Therefore, comparisons with MC or hybrid algorithms are not straightforward. The time steps chosen here are such that the time step error is divided homogeneously over the different modes.

The algorithm for a step in the Fourier-accelerated Langevin method is as follows:

ROUTINE LangStep(ψ)
  Calculate forces F_n in real space;
  FFT: $F_n \rightarrow F_k$;
  FFT: $\phi_n \rightarrow \phi_k$;
  Generate random forces R_k;
  Update ϕ_k using (15.103) with time steps (15.105);
  $\overline{\text{FFT}}$: $\phi_k \rightarrow \phi_n$;
END LangStep.

We have used 'FFT' for the forward transform (from real space to reciprocal space) and '$\overline{\text{FFT}}$' for the backward transform. The random forces R_k can be generated in two ways. The simplest way is to generate a set of random forces R_n on the real

space grid, and then Fourier transforming this set to the reciprocal grid. A more efficient way is to generate the forces directly on the Fourier grid. The forces R_k satisfy the following requirements. (i) $R_k = R^*_{-k}$, as a result of the R_n being real. (ii) The variance satisfies $\langle |R_k|^2 \rangle = \langle R_n^2 \rangle = 1$. Thus, for $k \neq -k$ (modulo $2\pi/L$), the real and imaginary part of the random force R_k both have width $1/\sqrt{2}$. If $k = -k$ (modulo $2\pi/L$) then the random force has a real part only, which should be drawn from a distribution with width 1.

15.6 Comparison of algorithms for scalar field theory

In the previous sections we have described seven different methods for simulating the scalar field theory on a lattice. We now present a comparison of the performance of the different methods. We have taken $m = 0.1$ and $g = 0.01$ as the bare parameters on a 16×16 lattice. The simulations were carried out on a standard workstation. The results should not be taken too seriously because different platforms and different, more efficient codings could give different results. Moreover, some methods can be parallelised more efficiently than others, which is important when doing large scale calculations (see Chapter 16). Finally, no real effort has been put into optimising the programs (except for standard optimisation at compile time), so the results should be interpreted as trends rather than as rigorous comparisons.

We give the CPU time needed for one simulation step and the correlation time, measured in simulation steps. The error in the run time is typically a few per cent, and that in the correlation time is typically between 5 and 10 per cent. For each method we include results for an 8×8 and a 16×16 lattice to show how the CPU time per step and the correlation time scale with the lattice size. All programs give the correct results for the renormalised mass and coupling constant, which have been presented before. The number of MC or MD steps in these simulations varied from 30 000 to 100 000, depending on the method used. In the Andersen method, we used 100 steps between momentum refreshing for $h = 0.05$ and 50 steps for $h = 0.1$. The time steps used in the hybrid algorithm were chosen such as to stabilise the acceptance rate at 70 per cent. In this algorithm, 10 steps were used between the updates. The time step $h = 0.2$ given for the Fourier-accelerated Langevin method is in fact the proportionality constant between h_k and the inverse propagator:

$$h_k = \frac{0.2}{\sqrt{K_{k,-k}}}. \tag{15.106}$$

From the table it is seen that for small lattices the heat bath and the hybrid methods are most efficient. For larger lattices, the multigrid and Fourier-accelerated

Table 15.3: Comparison between different methods for simulating the scalar quantum field theory on a lattice. The time units are only relative: no absolute run times should be deduced from them. The correlation time is measured in simulation steps (MD steps or MCS). For the methods with momentum refreshment, the correlation time is measured in MD steps. The overall efficiency in the last column is the inverse of (CPU time × correlation time). For 16×16 lattices this number has been multiplied by four.

Method	Described in section	Lattice size	Time constant h	Correlation time	CPU time	Overall efficiency
Metropolis	15.4.1	8		28	110	0.32
Metropolis	15.4.1	16		97	416	0.10
Heat bath	15.4.1	8		6.5	103	1.49
Heat bath	15.4.1	16		24	392	0.43
Andersen	15.4.3.1	8	0.05	170	32	0.063
Andersen	15.4.3.1	8	0.1	85	31	0.38
Andersen	15.4.3.1	16	0.1	110	124	0.29
Hybrid	15.4.3.2	8	0.365	15	45	1.48
Hybrid	15.4.3.2	16	0.22	60	175	0.38
Langevin	15.4.3.3	8	0.1	560	82	0.022
Langevin	15.4.3.3	16	0.1	2200	322	0.0056
Multigrid	15.5.4	8		1.5	1440	0.46
Multigrid	15.5.4	16		1.5	5750	0.46
Four/Lang	15.5.5	8	0.2	10	118	0.85
Four/Lang	15.5.5	16		10	523	0.76

Langevin methods take over, where the latter seems to be more efficient. However, its efficiency decreases logarithmically with size (multigrid remains constant) and comparisons of the values obtained with this method and MC simulations are always a bit hazardous, although these results may be very useful in themselves.

15.7 Gauge field theories

15.7.1 The electromagnetic Lagrangian

The scalar field theory is useful for some applications in particle physics and statistical mechanics – however, the fundamental theories describing elementary

particles have a more complicated structure. They include several kinds of particles, some of which are fermions. Intermediate particles act as 'messengers' through which other particles interact. It turns out that the action has a special kind of local symmetry, the so-called 'gauge symmetry'.

Global symmetries are very common in physics: rotational and translational symmetries play an important role in the solution of classical and quantum mechanical problems. Such symmetries are associated with a transformation (rotation, translation) of the full space, which leaves the action invariant. Local symmetries are operations which vary in space–time, and which leave the action invariant. You have probably met such a local symmetry: electrodynamics is the standard example of a system exhibiting a local gauge symmetry. The behaviour of electromagnetic fields in vacuum is described by an action defined in terms of the four-vector potential $A_\mu(x)$ (x is the space–time coordinate):[40]

$$S_{\mathrm{EM}} = \frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} \equiv \int d^4x\, \mathscr{L}_{\mathrm{EM}}(\partial_\mu A^\nu) \tag{15.107a}$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{15.107b}$$

$\mathscr{L}_{\mathrm{EM}} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ is the electromagnetic Lagrangian. The gauge symmetry of electrodynamics is a symmetry with respect to a particular class of space–time dependent shifts of the four-vector potential $A_\mu(x)$:

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\chi(x), \tag{15.108}$$

where $\chi(x)$ is some scalar function. It is easy to check that the action (15.107a) is indeed invariant under the gauge transformation (15.108). If sources j_μ are present [$j = (\rho, \mathbf{j})$ where ρ is the charge density, and $\mathbf{j}$ the current density], the action reads

$$S_{\mathrm{EM}} = \frac{1}{4}\int d^4x (F_{\mu\nu}F^{\mu\nu} + j_\mu A^\mu). \tag{15.109}$$

The Maxwell equations are found as the Euler-Lagrange equations for this action. The action is gauge invariant if the current is conserved, according to

$$\partial_\mu j^\mu(x) = 0. \tag{15.110}$$

A quantum theory for the electromagnetic field (in the absence of sources) is constructed proceeding in the standard way, by using the action (15.107a) in the path integral. If we fix the gauge, for example by setting $\partial_\mu A^\mu = 0$ (Lorentz gauge), the transition probability for going from an initial field configuration A_{i} at t_{i} to A_{f}

at t_f for imaginary times (we use Euclidean metric throughout this section) is given by

$$\langle A_f; t_f | A_i; t_i \rangle = \int [\mathscr{D} A_\mu] \exp\left[-\frac{1}{\hbar} \int_{t_i}^{t_f} dt\, L_{EM}(\partial_\mu A^\nu) \right] \tag{15.111}$$

where the path integral is over all vector potential fields which are compatible with the Lorentz gauge and with the initial and final vector potential fields at times t_i and t_f respectively. If we do not fix the gauge, this integral diverges badly, whereas for a particular choice of gauge, the integral converges.

Just as in the case of scalar fields, the excitations of the vector potential field are considered as particles. These particles are massless: they are the well-known *photons*. The electromagnetic field theory is exactly solvable: the photons do not interact, so we have a situation similar to the free field theory. The theory becomes more interesting when electrons and positrons are coupled to the field. These particles are described by vector fields $\psi(x)$ with $D = 2^{[d/2]}$ components for d-dimensional space–time ($[x]$ denotes the integer part of x), so $D = 4$ in four-dimensional space–time ($d = 4$). The first two components of the four-vector correspond to the spin-up and down states of the fermion (e.g. the electron) and the third and fourth components to the spin-up and down components of the anti-fermion (positron). The Euler-Lagrange equation for a fermion system interacting with an electromagnetic field is the famous *Dirac equation*:

$$\left[\gamma^\mu (\partial_\mu - ieA_\mu) + m \right] \psi(x) = 0. \tag{15.112}$$

The objects γ_μ are Hermitian $D \times D$ matrices obeying the anti-commutation relations:

$$[\gamma_\mu, \gamma_\nu]_+ = \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\delta_{\mu\nu} \tag{15.113}$$

(in Minkowski metric, $\delta_{\mu\nu}$ is to be replaced by $g_{\mu\nu}$). The Dirac equation is invariant under the gauge transformation (15.108) if it is accompanied by the following transformation of the ψ:

$$\psi(x) \to e^{ie\chi(x)} \psi(x). \tag{15.114}$$

The action from which the Dirac equation can be derived as the Euler-Lagrange equation is the famous quantum electrodynamics (QED) action:

$$S_{QED} = \int d^4x \Big[-\bar{\psi}(x)(\gamma^\mu \partial_\mu + m)\psi(x) + ieA_\mu(x)\bar{\psi}(x)\gamma^\mu \psi(x) - \frac{1}{4} F_{\mu\nu}(x) F^{\mu\nu}(x) \Big]. \tag{15.115}$$

Here, $\psi(x)$ and $\bar{\psi}(x)$ are independent fields. The Dirac equation corresponds to the Euler-Lagrange equation of this action with

$$\bar{\psi}(x) = \psi^\dagger(x)\gamma^0. \tag{15.116}$$

The QED action itself does not show the fermionic character of the ψ-field, which should however not disappear in the Lagrangian formulation. The point is that the ψ field is not an ordinary c-number field, but a so-called *Grassmann field*. Grassmann variables are anti-commuting numbers – Grassmann numbers a and b have the properties:

$$ab + ba = 0. \tag{15.117}$$

In particular, taking $a = b$, we see that $a^2 = 0$. We do not go into details concerning Grassmann algebra[4, 5, 41] but mention only the result of a Gaussian integration over Grassman variables. For a Gaussian integral over a vector $\boldsymbol{\psi}$ we have the following result for the components of $\boldsymbol{\psi}$ being ordinary commuting, or Grassmann anti-commuting numbers:

$$\int d\psi_1 \dots d\psi_N \exp(-\boldsymbol{\psi}^{\mathrm{T}} A \boldsymbol{\psi}) = \begin{cases} \sqrt{\frac{(2\pi)^N}{\det(A)}} & \text{commuting;} \\ \sqrt{\det(A)} & \text{anti-commuting.} \end{cases} \tag{15.118}$$

The matrix A is symmetric. In quantum field theories such as QED, we need a Gaussian integral over complex commuting and noncommuting variables, with the result:

$$\int d\psi_1 d\psi_1^* \dots d\psi_N d\psi_N^* \exp(-\boldsymbol{\psi}^{\dagger} A \boldsymbol{\psi}) = \begin{cases} (2\pi)^N / \det(A) & \text{commuting;} \\ \det(A) & \text{anti-commuting} \end{cases} \tag{15.119}$$

for a Hermitian matrix A. Fortunately the Lagrangian depends only quadratically on the fermionic fields, so only Gaussian integrals over Grassmann variables occur in the path integral.

15.7.2 Electromagnetism on a lattice – quenched compact QED

Physical quantities involving interactions between photons and electrons, such as scattering amplitudes, masses and effective interactions can be derived from the QED Lagrangian in a perturbative analysis. This leads to divergences similar to those mentioned in connection with scalar fields, and these divergences should be renormalised properly by choosing values for the bare coupling constant e and mass m occurring in the Lagrangian such that physical mass and coupling constant become finite – more precisely, they become equal to the experimental electron mass and the charge which occurs in the large-distance Coulomb law in three spatial dimensions:

$$V(r) = \frac{e^2}{4\pi\epsilon_0 r} \tag{15.120}$$

(for short distances, this formula is no longer valid as a result of quantum corrections).

Instead of following the perturbative route, we consider the discretisation of electrodynamics on a lattice (the Euclidean metric is most convenient for lattice calculations, so it is assumed throughout this section). This is less straightforward than in the scalar field case as a result of the greater complexity of the QED theory. We work in a space–time dimension $d = 4$. We first consider the discretisation of the photon gauge field and describe the inclusion of fermions below. An important requirement is that the gauge invariance should remain intact. Historically, Wegner's Ising lattice gauge theory[42] showed the way to the discretisation of continuum gauge theories. We now describe the lattice formulation for QED which was first given by Wilson[43] and then show that the continuum limit for strong coupling is the conventional electromagnetic gauge theory.

We introduce the following objects, living on the *links* μ of a square lattice with sites denoted by n:

$$U_\mu(n) = \exp\left[ieaA_\mu(n)\right] = \exp\left[i\theta_\mu(n)\right] \tag{15.121}$$

where we have defined the dimensionless scalar variables $\theta_\mu = eaA_\mu$. The action on the lattice is then written as a sum over all plaquettes, where each plaquette carries an action (see Figure 15.6):

$$S_{\text{plaquette}}(n;\mu\nu) = \text{Re}\left[1 - U_\mu(n)U_\nu(n+\mu)U_\mu^*(n+\mu+\nu)U_\nu^*(n+\nu)\right]. \tag{15.122}$$

Note that the effect of complex conjugation is a sign-reversal of the variable θ_μ. The

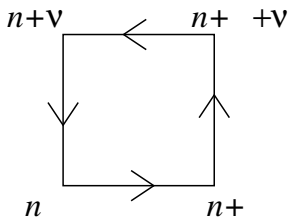

Figure 15.6: A lattice plaquette at site n with sides μ and ν used in (15.122).

U's are orientation-dependent: $U_\mu(n) = U_{-\mu}^\dagger(n)$. The constant 1 has been included in (15.122) to ensure that the total weight of a configuration with all θ-values being equal to zero vanishes. Note that the integration over θ is over a range 2π, so it does not diverge, in contrast to the original formulation, where the gauge must be

fixed in order to prevent the path integral on a finite lattice from becoming infinite [see also the remark after Eq. (15.111)]. The plaquette action can also be written as

$$S_{\text{plaquette}}(n;\mu\nu) = \mu\nu\left[1-\cos\left(\theta_{\mu\nu}(n)\right)\right] \tag{15.123}$$

where the argument of the cosine is the sum over the θ-variables around the plaquette as in Figure 15.6:

$$\theta_{\mu\nu}(n) = \theta_\mu(n) + \theta_\nu(n+\mu) - \theta_\mu(n+\mu+\nu) - \theta_\mu(n+\nu). \tag{15.124}$$

The total action

$$S_{\text{LQED}} = \sum_{n;\mu\nu} S_{\text{plaquette}}(n;\mu\nu) \tag{15.125}$$

occurs in the exponent of the time evolution operator or of the Boltzmann factor (for field theory in imaginary time). The partition function of the Euclidean field theory is

$$Z_{\text{LQED}}(\beta) = \int_0^{2\pi} \prod_{n,\mu} d\theta_\mu(n) \exp\left(-\beta S_{\text{LQED}}\right), \tag{15.126}$$

where the product $\prod_{n,\mu}$ is over all the links of the lattice. For low temperature (large β), only values of θ close to 0 (mod 2π) will contribute significantly to the integral. Expanding the cosine for small angles, we can extend the integrals to the entire real axis and obtain

$$S_{\text{LQED}}(\beta \text{ large}) = \sum_{n,\mu,\nu} \frac{1}{2}\left(\mu\nu\theta(n)\right)^2. \tag{15.127}$$

Using $\theta_\mu = eaA_\mu$ and the fact that the lattice constant a is small, we see that the action can be rewritten as

$$\beta S_{\text{LQED}} \approx \frac{\beta}{4}\int \frac{d^4x}{a^4}\left[a^4e^2F_{\mu\nu}(x)F^{\mu\nu}(x)\right], \tag{15.128}$$

where now the summation is over *all* $\mu\nu$, whereas in the sums above (over the plaquettes), μ and ν were restricted to positive directions. The $F_{\mu\nu}$ are defined in Eq. (15.107b). Taking $\beta = 1/e^2$ we recover the Maxwell Lagrangian:

$$\beta S_{\text{LQED}} \approx \frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} \tag{15.129}$$

in the continuum limit.

What are interesting objects to study? Physical quantities are gauge invariant, so we search for gauge-invariant correlation functions. Gauge invariance can be

formulated in the lattice model as an invariance under a transformation defined by a lattice function $\chi(n)$ which induces a shift in the $\theta_\mu(n)$:

$$\theta_\mu(n) \to \theta_\mu(n) + \frac{\chi(n+\mu) - \chi(n)}{a}. \tag{15.130}$$

This suggests that gauge-invariant correlation functions are defined in terms of a sum over θ_μ over a closed path: in that case a gauge transformation does not induce a change in the correlation function since the sum over the finite differences of the gauge function $\chi(n)$ over the path will always vanish. Furthermore, as correlation functions usually contain products of variables at different sites, we consider the so-called *Wilson loop* correlation function:

$$W(C) = \left\langle \prod_{n,\mu \epsilon C} e^{i\theta_\mu(n)} \right\rangle, \tag{15.131}$$

where the product is over all links n, μ between site n and its neighbour $n+\mu$ lying on the closed loop C – see Figure 15.7.[43]

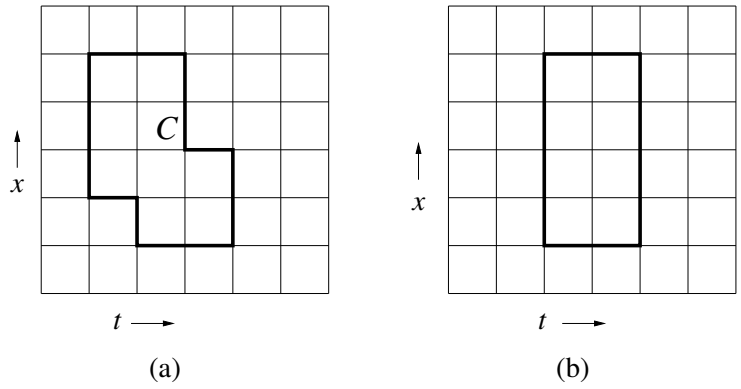

Figure 15.7: The Wilson loop on a two-dimensional square lattice. (a) shows a general Wilson loop and (b) shows a two-fermion loop in a gauge field theory with infinite mass fermions.

The Wilson loop correlation function has a physical interpretation. Suppose we create at some time t_i a fermion-antifermion pair, which remains in existence at fixed positions up to some time t_f, at which the pair is annihilated again. Without derivation we identify the partition function of the gauge field in the presence of the fermion-antifermion pair with the Wilson loop correlation function in Figure 15.7(b) times the vacuum partition function – for a detailed derivation see

refs. 6, 44, 45. Now let us stretch the loop in the time direction, $T = t_f - t_i \to \infty$. The effective interaction between two electrons at a distance R is given by the difference between the ground state energy in the presence of the fermion-antifermion pair (which we denote by 2f) and the ground state energy of the vacuum:

$$V(R) = \left\langle \psi_G^{(2f)} \middle| H \middle| \psi_G^{(2f)} \right\rangle_{2f} - \left\langle \psi_G^{(vac)} \middle| H \middle| \psi_G^{(vac)} \right\rangle_{vac} . \tag{15.132}$$

This expression can however be evaluated straightforwardly in the Lagrangian picture. We have

$$e^{-TV(R)} = \frac{Z(C)}{Z} = W(C) \tag{15.133}$$

where C is the rectangular contour of size T (time direction) and R (space direction); $Z(C)$ is the partition function evaluated in the presence of the Wilson loop, and Z is the vacuum partition function. Note that the fact that T is taken large, guarantees that the ground state of the Hamiltonian is projected out in the simulation.

By varying the coupling β, different results for the value of the Wilson loop correlation function are found. For large loops we have either the 'area law':

$$W(C) = \exp\big(\text{Const.} \times \text{Area within loop}\big), \tag{15.134}$$

or the 'perimeter law' :

$$W(C) = \exp\big(\text{Const.} \times \text{Perimeter of loop}\big), \tag{15.135}$$

with additional short-corrections. Let us consider the area law. In that case we find $V(R) \sim R$, which means that the two particles cannot be separated: pulling them infinitely far apart requires an infinite amount of energy. We say that the particles are *confined*. On the other hand, the perimeter law says that $V(R)$ is a constant (it is dominated by the vertical parts of length T), up to corrections decaying to zero for large R (for the confined case, these corrections can be neglected). For $d = 4$ one finds after working out the dominant correction term $V(R) \sim \text{Const.} - e^2/R$, i.e. Coulomb's law.[44] We see that the lattice gauge theory incorporates two different kinds of gauge interactions: confined particles and electrodynamics. The analysis in which the fermions are kept at fixed positions corresponds to the fermions having an infinite mass. It is also possible to allow for motion of the fermions by allowing the loops of arbitrary shape, introducing gamma-matrices in the resulting action. The procedure in which the fermions are kept at fixed positions is called 'quenched QED' – in quenched QED, vacuum polarisation effects (caused by the fact that photons can create electron-positron pairs) are not included.

We know that electrodynamics does not confine electrons: the lattice gauge theory in four dimensions has two phases, a low-temperature phase in which

the interactions are those of electrodynamics, and a high-temperature phase in which the particles are confined[46] ('temperature' is inversely proportional to the coupling constant β). The continuum limit of electrodynamics is described by the low-temperature phase of the theory. Why have people been interested in putting QED on a grid? After all, perturbation theory works very well for QED, and the lattice theory gives us an extra phase which does not correspond to reality (for QED). The motivation for studying lattice gauge theories was precisely this latter phase: we know that quarks, the particles which are believed to be the constituents of mesons and hadrons, are confined: an isolated quark has never been observed. Lattice gauge theory provides a mechanism for confinement! Does this mean that quarks are part of the same gauge theory as QED, but corresponding to the high-temperature phase? No: there are reasons to assume that a quark theory has a more complex structure than QED, and moreover, experiment has shown that the interaction between quarks vanishes when they come close together, in sharp contrast with the confining phase of electrodynamics in which the interaction energy increases linearly with distance. The high-temperature phase of the gauge theory considered so far is always confining, so this does not include this short-distance decay of the interaction, commonly called 'asymptotic freedom'.[6, 44, 45, 47, 48] We shall study the more complex gauge theory which is believed to describe quarks later – this theory is called 'quantum chromodynamics' (QCD).

The lattice version of quantum electrodynamics using variables θ_μ ranging from 0 to 2π is often called $U(1)$ lattice gauge theory because the angle θ_μ parametrises the unit circle, which in group theory is called $U(1)$. Another name for this field theory is 'compact QED' because the values assumed by the variable θ_μ form a compact set, as opposed to the noncompact A_μ field of continuum QED. Compact QED can be formulated in any dimension, and in the next section we discuss an example in 1 space + 1 time dimension.

15.7.3 A lattice QED simulation

We describe a QED lattice simulation for determining the inter-fermion potential. We do this by determining the Wilson loop correlation function described in the previous section. Only the gauge field is included in the theory – the fermions have a fixed position, and the photons exchanged between the two cannot generate fermion–antifermion pairs (vacuum polarisation). This is equivalent to assigning an infinite mass to the fermions. We use a square lattice with periodic boundary conditions.

We consider the 1+1-dimensional case. This is not a very interesting theory by itself – it describes confined fermions, as the Coulomb potential in one spatial

dimension is confining:

$$V(x) \sim |x|, \tag{15.136}$$

but we treat it here because it is simple and useful for illustrating the method. The theory can be solved exactly (see problem 15.6):[44] the result is that the Wilson loop correlation function satisfies the area law:

$$W = \exp(-\alpha A) \tag{15.137}$$

(A is the area enclosed within the loop) with the proportionality constant α given in terms of the modified Bessel functions I_n:

$$\alpha = -\ln\left[\frac{I_1(\beta)}{I_0(\beta)}\right]; \tag{15.138}$$

β is the coupling constant (inverse temperature). In fact the area law holds only for loops much smaller than the system size; deviations from this law occur when the linear size of the loop approaches half the system size.

The system can be simulated straightforwardly using the Metropolis algorithm, but we shall use the heat bath algorithm because of its greater efficiency. We want the coefficient α to be not too large, as large values of α cause $W(C)$ to decay very rapidly with size, so that it cannot be distinguished from the simulation noise for loops of a few lattice constants. From (15.138) we see that β must be large in that case – we shall use $\beta = 10$. This causes the probability distribution $P(\theta_\mu)$ for some θ_μ, embedded in a particular, fixed configuration of θ_μ on neighbouring links, to be sharply peaked. Therefore it is not recommended to take θ_μ random between 0 and 2π and then accept with probability $P(\theta_\mu)$ and retry otherwise, as in this approach most trial values would end up being rejected. We shall therefore first generate a trial value for θ_μ according to a Gaussian probability distribution.

The distribution $P(\theta)$ has the form

$$P(\theta) = \exp\left\{-\beta\left[\cos(\theta-\theta_1)+\cos(\theta-\theta_2)\right]\right\} \tag{15.139}$$

where θ_1 and θ_2 are fixed – they depend on the θ-values on the remaining links of the plaquettes containing θ. The sum of the two cosines can be rewritten as

$$\cos(\theta-\theta_1)+\cos(\theta-\theta_2) = 2\cos\left(\frac{\theta_1-\theta_2}{2}\right)\cos\left(\theta-\frac{\theta_1+\theta_2}{2}\right). \tag{15.140}$$

We define

$$\tilde{\beta} = 2\beta\cos\left(\frac{\theta_1-\theta_2}{2}\right) \text{ and} \tag{15.141a}$$

$$\phi = \theta - \frac{\theta_1+\theta_2}{2} \tag{15.141b}$$

so that our task is now to generate an angle ϕ with a distribution $\exp(-\tilde{\beta}\cos\phi)$. We distinguish between two cases: (i) $\tilde{\beta} > 0$. In that case the maximum of the distribution occurs at $\phi = \pi$. A Gaussian distribution centred at π and with width $\sigma = \pi/(2\sqrt{\tilde{\beta}})$ and amplitude $\exp(\tilde{\beta})$ is always close to the desired distribution. The Gaussian random numbers must be restricted to the interval $[-\pi, \pi]$. Therefore the algorithm becomes:

REPEAT
  REPEAT
    Generate a Gaussian random variable $-r$ with width 1;
    $\phi = \sigma r$;
  UNTIL $-\pi \le \phi \le \pi$;
  $\phi \to \phi + \pi$;
  Accept this trial value with probability
    $\exp\{-\tilde{\beta}[\cos\phi + 1 - (\phi - \pi)^2/(2\sigma^2)]\}$;
UNTIL Accepted;
$\theta = \phi + \frac{\theta_1 + \theta_2}{2}$.

(ii) If $\tilde{\beta} < 0$ then the distribution is centred around $\phi = 0$. In that case, we do not shift the Gaussian random variable over π. The reader is invited to work out the analogue of the algorithm for case (i).

In the simulation we calculate the averages of square Wilson loops, given in Eq. (15.131) (it should be emphasised that for the area law it is not required to have $T \gg R$). This is done by performing a loop over all lattice sites and calculating the sum of the θ_μ over the square loop with lower left corner at the current site. The expectation values for different square sizes can be calculated in a single simulation. Figure 15.8 shows the average value of the Wilson loop correlation functions as a function of the area enclosed by the loop for a 16×16 and a 32×32 lattice. The drawn straight line is the theoretical curve with slope α as in (15.138). From this figure it is seen that the area law is satisfied well for loops which are small with respect to half the lattice size. By implementing free boundary conditions, the theoretical curve can be matched exactly, but this requires a little more bookkeeping.

15.7.4 Including dynamical fermions

In real problems studied by particle physicists, fermions are to be included into the lattice action, and not in a quenched fashion as done in the previous section. In this section we focus on dynamical fermions, which cause two problems. First of all, a straightforward discretisation of the fermion action leads to 2^d species of uncoupled fermions to be included into the problem (in d space–time dimensions)

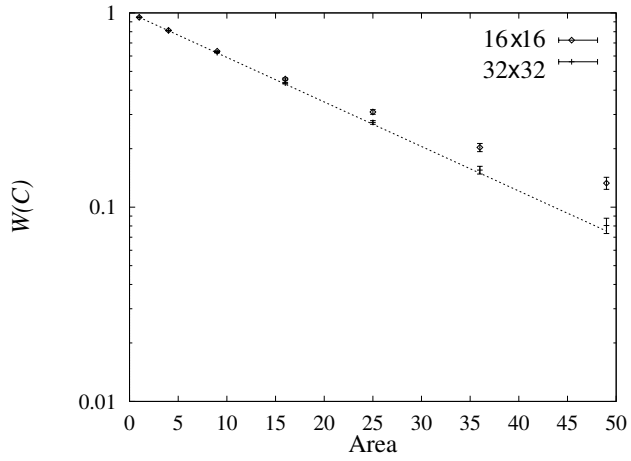

Figure 15.8: The Wilson loop correlation function as a function of the enclosed area for 1+1 dimensional lattice QED. Note that the vertical scale is logarithmic, so that the straight line is compatible with the area law. The values were determined in a heat bath simulation using 40 000 updates (first 2000 rejected).

instead of the desired number of fermion species ('flavours'). Second, we have not yet discussed the problem of including the fermion character in a path integral simulation. We first consider the 'fermion doubling problem' and then sketch how simulations can actually be carried out for theories including fermions.

15.7.4.1 Fermions on a lattice

When calculating the path integral for free fermions for which the Lagrangian is quadratic in the fermion fields, the following Gaussian Grassmann integral must be evaluated:

$$\int \left[\mathscr{D}\psi\mathscr{D}\bar{\psi}\right] e^{-\bar{\psi}M\psi}, \tag{15.142}$$

where the kernel M is given as

$$M = m + \gamma^{\mu}\partial_{\mu}. \tag{15.143}$$

The expression in the exponent is short-hand for an integral over the space–time coordinates. Discretising the theory on the lattice and Fourier transforming the fields and M we find that the latter becomes diagonal:

$$M_{k,k'} = \left(m + \frac{i}{a}\sum_{\mu}\gamma_{\mu}\sin(k_{\mu}a)\right)\delta(k+k'). \tag{15.144}$$

The lattice version of M is therefore the Fourier transform of this function.

There is a problem with this propagator. The continuum limit singles out only the minima of the sine as a result of the factor $1/a$ in front of it. These are found not only near $k = 0$ but also near $ka = (\pm\pi, 0, 0, 0)$ (in four dimensions) etcetera, because of the sine function having zeroes at 0 and π. This causes the occurrence of two different species of fermions per dimension, adding up to 16 species for four-dimensional space–time. It turns out that this degeneracy can be lifted only at the expense of breaking the so-called 'chiral symmetry' for massless fermions.[49] Chiral symmetry is a particular symmetry which is present in the Dirac equation (and action) for massless particles. Suppose chiral symmetry is present in the lattice version of the action. This symmetry forbids a mass term to be present, and the renormalised theory should therefore have $m = 0$. A lattice action which violates chiral symmetry might generate massive fermions under renormalisation.

One could ignore the doubling problem and live with the fact that the theory now contains 2^d different species of fermions. However, it is also possible to lift up the unwanted parts of the propagator by adding a term proportional to $1 - \cos(ak_\mu)$ to it, which for k near 0 does not affect the original propagator to lowest order, but which lifts the parts for $k_\mu a = \pm\pi$ such that they are no longer picked up in the contiuum. The method is referred to as the *Wilson fermion method*. The resulting propagator is

$$M_k = m + \frac{i}{a}\sum_\mu \gamma_\mu \sin(ak_\mu) + \frac{r}{a}\sum_\mu \left[1 - \cos(ak_\mu)\right]. \tag{15.145}$$

This form is very convenient because it requires only a minor adaptation of a program with the original version of the propagator. In real space, and taking the lattice constant a equal to 1, the Wilson propagator reads in d space–time dimensions:

$$M_{nl} = (m + 4r)\delta_{nl} - \frac{1}{2}\sum_\mu \left[(r + \gamma_\mu)\delta_{l,n+\mu} + (r - \gamma_\mu)\delta_{l,n-\mu}\right]. \tag{15.146}$$

The disadvantage of this solution is that the extra terms destroy any chiral symmetry, which is perhaps a bit too drastic.

In a more complicated method half of the unwanted states are removed by doubling the period, so that the Brillouin zone of the lattice is cut off at $\pi/(2a)$ instead of π/a. This is done by putting different species of fermions on alternating sites of the lattice. Although this removes the unwanted fermions, it introduces new fermions which live on alternate sites of the lattice. The resulting method is called the *staggered fermion method*. The staggered fermion method respects chiral symmetry discussed above and is therefore a better option than the Wilson fermion

method. It is, however, more complicated than the Wilson fermion method and we refrain from a discussion here, but refer to the original literature[50, 51] and later reviews.[6, 45, 52]

15.7.4.2 Algorithms for dynamical fermions

If we want to include dynamical rather than quenched fermions into our lattice field theory, we must generate configurations of anti-commuting fermion fields. As it is not clear how to do this directly and as this may cause negative probabilities, various alternatives using results for Gaussian integrals over Grassmann variables (see Section 15.7.1) have been developed. We shall explain a few algorithms for an action consisting of a bosonic part, S_{Boson}, defined in terms of the boson field, $A(x)$, coupled to the fermion field, $\psi, \overline{\psi}$, via the fermion kernel, $M(A)$:

$$S = S_{\text{Boson}}(A) + \int d^d x \bar{\psi}(x) M(A) \psi(x). \tag{15.147}$$

The QED Lagrangian in (15.115) has this form.

Integrating out the fermion part of the path integral using (15.119) leads to a path integral defined entirely in terms of bosons:

$$\begin{aligned}\int [\mathscr{D}A][\mathscr{D}\overline{\psi}][\mathscr{D}\psi] e^{-[S_{\text{Boson}}(A)+\overline{\psi}M(A)\psi]} &= \int [\mathscr{D}A] \det[M(A)] e^{-\mathscr{L}_{\text{Boson}}(A)} \\ &= \int [\mathscr{D}A] e^{-\mathscr{L}_{\text{Boson}}(A)+\ln[\det(M(A))]}\end{aligned} \tag{15.148}$$

(the inverse temperature β is included in the action). Although the determinant of $M(A)$ is real and usually positive, $M(A)$ is not necessarily a positive definite Hermitian matrix (a positive definite matrix has real and positive eigenvalues). It is therefore sometimes useful to consider the matrix

$$W(A) = M^\dagger(A) M(A), \tag{15.149}$$

in terms of which the path integral can be written as

$$\int [\mathscr{D}A] e^{-\mathscr{L}_{\text{Boson}}(A)+\frac{1}{2}\ln[\det(W(A))]}. \tag{15.150}$$

Now suppose that we want to perform a Metropolis update of the A-field. The acceptance probability for a trial change $A \to A'$ is

$$\begin{aligned}P_{\text{Accept}}(A \to A') &= e^{-S_{\text{Boson}}(A')+S_{\text{Boson}}(A)} \frac{\det[M(A')]}{\det[M(A)]} \\ &= e^{-S_{\text{Boson}}(A')+S_{\text{Boson}}(A)} \sqrt{\frac{\det[W(A')]}{\det[W(A)]}}.\end{aligned} \tag{15.151}$$

This is very expensive to evaluate since we must calculate the determinant of the very large matrix $M(A)$ [or $W(A)$] at every step. A clever alternative follows from the observation that if the field A changes on one site only (as is usually the case), very few elements of the matrix $M(A)$ change, which allows us to perform the calculation more efficiently.[53] Another interesting suggestion is that of Bhanot *et al.*[54] who propose to evaluate the fraction of the determinants as follows:

$$\frac{\det[W(A')]}{\det[W(A)]} = \frac{\int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger W(A)\phi}}{\int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger W(A')\phi}}, \tag{15.152}$$

where ϕ is a boson field for which we can use the algorithms given earlier in this chapter. Defining $\Delta W = W(A') - W(A)$, we can express the ratio in terms of an expectation value:

$$\frac{\det[W(A')]}{\det[W(A)]} = \left\langle \exp(\phi^\dagger \Delta W \phi) \right\rangle_{A'} = 1 / \left\langle \exp(-\phi^\dagger \Delta W \phi) \right\rangle_{A}. \tag{15.153}$$

It is now possible to calculate this average by updating the field ϕ in a heat bath algorithm. As the matrix $M(A)$ is local (it couples only nearest neighbours), $W(A)$ is local as well (it couples up to next nearest neighbours). Therefore the heat bath algorithm can be carried out efficiently (it should be possible to apply the SOR method to this method). Each time we change the field A, the matrix $W(A)$ changes and a few heat bath sweeps for the field ϕ have to be carried out. The value of the fraction of the determinants is determined as the geometrical average of the two estimators given in Eq. (15.153).

The most efficient algorithms for dynamical fermions combine a molecular dynamics method for the boson fields with a Monte Carlo approach for the fermionic part of the action. We describe two of these here. The first one is a Langevin approach, proposed by Batrouni *et al.*,[17] and suitable for Fourier acceleration. It is based on two observations: first, $\det(M)$ can be written as $\exp[\mathrm{Tr}\ln(M)]$, and second, if ξ_n is a complex Gaussian random field on the lattice, so that

$$\left\langle \xi_l^\dagger \xi_n \right\rangle = \delta_{nl} \tag{15.154}$$

(the brackets $\langle\rangle$ denote an average over the realisations of the Gaussian random generator), then the trace of any matrix K can be written as

$$\mathrm{Tr}(K) = \sum_{nl} \left\langle \xi_n^\dagger K_{nl} \xi_l \right\rangle. \tag{15.155}$$

In the Langevin approach, the force is given by the derivative of the action with respect to the boson field. In the presence of fermions, the action reads [see Eq. (15.148)]:

$$S = S_{\mathrm{Boson}} - \mathrm{Tr}\ln[M(A)]. \tag{15.156}$$

Therefore the derivative has the form

$$\frac{\partial S(A)}{\partial A_n} = \frac{\partial S_{\text{Boson}}(A)}{\partial A_n} - \text{Tr}\left[M^{-1}(A)\frac{\partial M(A)}{\partial A_n}\right]. \tag{15.157}$$

To evaluate the trace, we make use of the auxiliary field ξ:

$$\text{Tr}\left[M^{-1}(A)\frac{\partial M(A)}{\partial A_n}\right] = \left\langle \xi^\dagger M^{-1}(A)\frac{\partial M(A)}{\partial A_n}\xi \right\rangle \tag{15.158}$$

$$= \sum_{ijl}\left\langle \xi_i^* M^{-1}_{ij}(A)\left[\frac{\partial M(A)}{\partial A_n}\right]_{jl}\xi_l \right\rangle. \tag{15.159}$$

In the Langevin equation we do not calculate the average over the ξ by generating many random fields for each step, but instead we generate a single random Gaussian vector ξ at every MD step, and evaluate the terms in angular brackets in (15.158) only for this configuration. Below we justify this simplification. The MD step reads therefore [see Eq. (15.103)]:

$$A_n(t+h) = A_n(t) + \frac{h^2}{2}\left[-\frac{\partial S_{\text{Boson}}(A)}{\partial A_n} + \xi^\dagger M^{-1}(A)\frac{\partial M(A)}{\partial A_n}\xi\right] + h\eta_n. \tag{15.160}$$

The A-fields occurring between the square brackets are evaluated at time t. To evaluate the second term in the square brackets we must find the vector ψ satisfying

$$M(A)\psi = \xi, \tag{15.161}$$

so that the algorithm reads

$$A_n(t+h) = A(t) + \frac{h^2}{2}\left[-\frac{\partial S_{\text{Boson}}(A)}{\partial A_n} + \psi^\dagger\frac{\partial M(A)}{\partial A_n}\xi\right] + h\eta_n. \tag{15.162}$$

Finding the vector ψ is time-consuming. Use is made of the sparseness of the matrix $M(A)$ in order to speed up the calculation.[†] Note that this calculation is done only once per time step in which the full boson field is updated.

In the Langevin equation we generate a set of configurations which occur with a probability distribution given by the action (or rather an approximation to it because of time discretisation). If we evaluate the average distribution with respect to the random noise fields η and ξ, the average over the ξ-field gives us back the trace via equation (15.155), therefore we were justified in replacing the average over the noise field by the value for the actual noise field. It must be noted that Fourier acceleration is implemented straightforwardly in this fermion method: after the

[†]The conjugate gradient method (Section A.8.1.2) is applied to this matrix problem.

force is evaluated with the noise field ξ, it is Fourier transformed, and the leap-frog integration proceeds as described in Section 15.5.5.

Finally we describe a combination of MD and MC methods[55] which can be formulated within the hybrid method of Duane *et al.*;[18] see also Section 15.4.3.2.

A first idea is to replace the determinant by a path integral over an auxiliary boson field:

$$\det[M(A)] = \int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger M^{-1}(A).\phi}. \tag{15.163}$$

We want to generate samples of the auxiliary field ϕ with the appropriate weight. Equation (15.163) is, however, a somewhat problematic expression as it involves the inverse of a matrix which moreover is not Hermitian. If we have an even number of fermion flavours, we can group the fermion fields into pairs, and each pair yields a factor $\det[M(A)]^2$ which can be written as

$$\det[M(A)]^2 = \int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger [M(A)^\dagger M(A)]^{-1}\phi} = \int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger [W(A)]^{-1}\phi} \tag{15.164}$$

with $W(A)$ defined in (15.149). Note that we need an even number of fermion flavours here, because we cannot simply replace the matrix $W(A)$ by its square root in the following algorithm (see also the beginning of this subsection). This partition function is much more convenient than (15.163) for generating MC configurations of the field ϕ. This is done by an exact heat bath algorithm, in which a Gaussian random field ξ_n is generated, and the field ϕ is then found as

$$M(A)\xi = \phi. \tag{15.165}$$

For staggered fermions (see Section 15.7.4.1) it turns out that $W(A) = M(A)^\dagger M(A)$ couples only even sites with even sites, or odd sites with odd sites.[56, 57] Therefore the matrix $W(A)$ factorises into an even-even (ee) and an odd-odd one (oo), so that we can write

$$\det[W(A)] = \det[W(A)_{\mathrm{ee}}] \det[W(A)_{\mathrm{oo}}]. \tag{15.166}$$

The matrices $W(A)_{\mathrm{ee}}$ and $W(A)_{\mathrm{oo}}$ are identical – therefore we have:

$$\det[M(A)] = \int [\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-\phi^\dagger [W_{\mathrm{ee}}(A)]^{-1}\phi}. \tag{15.167}$$

The matrix W_{ee} is Hermitian and positive; it can be written as $W_{\mathrm{ee}} = M^\dagger_{\mathrm{eo}}(A) M_{\mathrm{oe}}(A)$, where the two partial matrices M_{eo} and M_{oe} are again identical, so we can use the heat bath algorithm as described, with $M(A)$ replaced by $M_{\mathrm{eo}}(A)$.

The full path integral contains only integrations over boson fields:

$$Z = \int [\mathscr{D}A][\mathscr{D}\phi][\mathscr{D}\phi^*] e^{-S_{\text{Boson}}(A) - \phi^\dagger W^{-1}(A)\phi} \tag{15.168}$$

where subscripts ee for W should be read in the case of staggered fermions. We want to formulate a molecular dynamics algorithm for the boson field A, but generate the auxiliary field configurations ϕ with an MC technique. This procedure is justified because the MD trajectory between an acceptance/rejection decision is reversible, and the acceptance/rejection step ensures detailed balance.

We assign momenta to the boson field A only:

$$Z = \int [\mathscr{D}A][\mathscr{D}\phi][\mathscr{D}\phi^*][\mathscr{D}P] e^{-\frac{1}{2}\sum_n P_n^2(x) - S_{\text{Boson}}(A) - \phi^\dagger W^{-1}(A)\phi}. \tag{15.169}$$

The equations of motion for the field A and its conjugate momentum P are then given by

$$\dot{A}_n = P_n; \tag{15.170a}$$

$$\dot{P}_n = -\frac{\partial S_{\text{Boson}}}{\partial A_n} - \sum_{lm} \phi_l^\dagger \frac{\partial \left[W_{lm}^{-1}(A)\right]}{\partial A_n} \phi_m. \tag{15.170b}$$

The difficult part is the second equation which involves the derivative of the inverse of $W(A)$. The key observation is now that

$$\frac{\partial W^{-1}(A)}{\partial A_n} = W^{-1}(A) \frac{\partial W(A)}{\partial A_n} W^{-1}(A), \tag{15.171}$$

so that we need the vector η with

$$W(A)\eta = \phi. \tag{15.172}$$

This can be found using a suitable sparse matrix algorithm. Using this η-field, the equation of motion for P simply reads:

$$\dot{P}_n = -\frac{\partial S_{\text{Boson}}}{\partial A_n} - \eta^\dagger \frac{\partial (M^\dagger M)}{\partial A_n} \eta. \tag{15.173}$$

Summarising, a molecular dynamics update consists of the following steps:

ROUTINE MDStep
  Generate a Gaussian random configuration ξ;
  Calculate $\phi = M(A)\xi$;
  Calculate η from $(M^\dagger M)\eta = \phi$;

Update the boson field A and its conjugate momentum field P using
$P(t+h/2) = P(t-h/2) - h\left[\frac{\partial S_{\text{Boson}}}{\partial A_n} + \eta\frac{\partial (M^{\mathrm{T}}M)}{\partial A_n}\eta\right]$ and
$A(t+h) = A(t) + hP(t+h/2)$.
END MDStep

We see that in both the Langevin and the hybrid method, the most time-consuming step is the calculation of a (sparse) matrix equation at each field update step (in the above algorithm this is the step in the third line).

15.7.5 Non-abelian gauge fields – quantum chromodynamics

QED is the theory for charged fermions interacting through photons, which are described by a real-valued vector gauge field A_μ. Weak and strong interactions are described by similar, but more complicated theories. A difference between these theories and QED is that the commuting complex phase factors $U_\mu(n)$ of QED are replaced by noncommuting matrices, members of the group SU(2) (for the weak interaction) or SU(3) (strong interaction). Furthermore, in quantum chromodynamics (QCD), the SU(3) gauge theory for strong interactions, more than one fermion flavour must be included. In this section we focus on QCD, where the fermions are the *quarks*, the building blocks of mesons and hadrons, held together by the gauge particles, called *gluons*. The latter are the QCD analogue of photons in QED.

Quarks occur in different species, or 'flavours' ('up', 'down', 'strange',...); for each species we need a fermion field. In addition to the flavour quantum number, each quark carries an additional colour degree of freedom: red, green or blue. Quarks form triplets of the three colours (hadrons, such as protons and neutrons), or doublets consisting of colour-anticolour (mesons) – they are always observed in colourless combinations. Quarks can change colour through the so-called *strong interactions*. The gluons are the intermediary particles of these interactions – they are described by a gauge field of the SU(3) group (see below). The gluons are massless, just as the photons in QED.

The U(1) variables of QED were parametrised by a single compact variable θ [$U = \exp(i\theta)$]. In QCD these variables are replaced by SU(3) matrices. These matrices are parametrised by eight numbers – they correspond to eight gluon fields A^a_μ, $a = 1,\ldots,8$ (gluons are insensitive to flavour). The gluons are massless, just as the photons, because inclusion of a mass term $m^2 A_\mu A^\mu$, analogous to that of the scalar field, destroys the required gauge invariance.

Experimentally, quarks are found to have almost no interaction at short separation, but when the quarks are pulled apart their interaction energy becomes linear with the separation, so that it is impossible to isolate one quark. The colour

interaction carried by the gluon fields is held responsible for this behaviour. There exists furthermore an intermediate regime, where the interaction is Coulomb-like.

The fact that the interaction vanishes at short distances is called 'asymptotic freedom'. It is possible to analyse the behaviour of quarks and gluons in the short-distance/small coupling limit by perturbation theory, which predicts indeed asymptotic freedom.[47, 58, 59] The renormalised coupling constant increases with increasing distance, and it is this coupling constant which is used as the perturbative parameter. At length scales of about 1 fm the coupling constant becomes too large, and the perturbative expansion breaks down. This is the scale of hadron physics. The break-down of perturbation theory is the reason why people want to study SU(3) gauge field theory on a computer, as this allows for a nonperturbative treatment of the quantum field theory. The lattice formulation has an additional advantage. If we want to study the time evolution of a hadron, we should specify the hadron state as the initial state. But the hadron state is very complicated! If we take the lattice size in the time direction large enough, the system will find the hadron state 'by itself' because that is the ground state, so that this problem does not occur.

The QCD action has the following form ($i = 1,2,3$ denotes the colour degree of freedom of the quarks, f the flavour):

$$S_{\mathrm{QCD}} = \int d^4x \left\{ \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \sum_f \sum_{ij} i\overline{\psi}^i_f \gamma^\mu \left(\delta_{ij}\partial_\mu + ig\frac{A^a_\mu}{2}\lambda^a_{ij} \right) \psi^j_f + \sum_f \sum_i m_f \overline{\psi}^i_f \psi^i_f \right\}. \quad (15.174)$$

The matrices λ^a are the eight generators of the group SU(3) [they are the Gell-Mann matrices, the analogue for SU(3) of the Pauli matrices for SU(2)], satisfying

$$\mathrm{Tr}(\lambda_a \lambda_b) = \delta_{a,b}. \quad (15.175)$$

The m_f are the quark masses, and $F^a_{\mu\nu}$ is more complicated than its QED counterpart:

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu; \quad (15.176)$$

the constants f^{abc} are the structure constants of $SU(3)$, defined by

$$[\lambda^a, \lambda^b] = 2i \sum_c f^{abc} \lambda^c. \quad (15.177)$$

The parameter g is the coupling constant of the theory; it plays the role of the charge in QED. A new feature of this action is that the f^{abc}-term in (15.176) introduces

interactions between the gluons, in striking contrast with QED where the photons do not interact. This opens the possibility to have massive gluon bound states, the so-called 'glueballs'.

When we regularised QED on the lattice, we replaced the gauge field A_μ by variables $U_\mu(n) = e^{ieA_\mu(n)}$ living on a link from site n along the direction given by μ. For QCD we follow a similar procedure: we put SU(3) matrices $U_\mu(n)$ on the links. They are defined as

$$U_\mu(n) = \exp\left(ig\sum_a A_\mu^a \lambda^a/2\right). \tag{15.178}$$

The lattice action is now constructed in terms of these objects. The gauge part of the action becomes

$$S_{\text{Gauge}} = \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \to -\frac{1}{g^2}\text{Tr}\left[U_\mu(n)U_\nu(n+\mu)U^\dagger_\mu(n+\nu)U^\dagger_\nu(n) + \text{Hermitian conjugate}\right]. \tag{15.179}$$

The quark part of the action, which includes the coupling with the gluons, reads in the case of Wilson fermions (see above):

$$S_{\text{Fermions}} = \sum_n (m+4r)\bar\psi(n)\psi(n) - \sum_{n,\mu}\left[\bar\psi(n)(r-\gamma^\mu)U_\mu(n)\psi(n+\mu) + \bar\psi(n+\mu)(r+\gamma^\mu)U^\dagger_\mu(n)\psi(n)\right]. \tag{15.180}$$

An extensive discussion of this regularisation, including a demonstration that its continuum limit reduces to the continuum action (15.174) can be found for example in Rothe's book.[45] The lattice QCD action

$$S_{\text{LQCD}} = S_{\text{Gauge}} + S_{\text{Fermions}} \tag{15.181}$$

can now be simulated straightforwardly on the computer, although it is certainly complicated. We shall not describe the procedure in detail. In the previous sections of this chapter we have described all the necessary elements, except for updating the gauge field, which is now a bit different because we are dealing with matrices as stochastic variables as opposed to numbers. Below we shall return to this point.

Simulating QCD on a four-dimensional lattice requires a lot of computer time and memory. A problem is that the lattice must be rather large. To see this, let us return to the simpler problem of quenched QCD, where the quarks have infinite mass so that they do not move; furthermore there is no vacuum polarisation in that case. The Wilson loop correlation function, now defined as

$$W(C) = \text{Tr}\prod_{(n,\mu)\epsilon C} U_\mu(n), \tag{15.182}$$

where the product is to be evaluated in a *path-ordered* fashion, i.e. the matrices must be multiplied in the order in which they are encountered when running along the loop. This is different from QED and reflects the fact that the U's are noncommuting matrices rather than complex numbers. This correlation function gives us the quenched inter-quark potential in the same way as in QED. In this approximation, perturbative renormalisation theory can be used to find an expansion for the potential at short distances in the coupling constant, g, with the result:

$$V(R,g,a) = \frac{C}{4\pi R}\left[g^2 + \frac{22}{16\pi^2}g^4 \ln\frac{R}{a} + \mathcal{O}(g^6)\right]. \tag{15.183}$$

Here C is a constant. We see that the coefficient of the second term increases for large R, rendering the perturbative expansion suspect, as mentioned before. The general form of this expression is

$$V(R,g,a) = \alpha(R)/R, \tag{15.184}$$

in other words, a 'screened Coulomb' interaction. Equation (15.183) can be combined with the requirement that the potential should be independent of the renormalisation cut-off a

$$a\frac{dV(r,g,a)}{da} = 0 \tag{15.185}$$

to find a relation between the coupling constant g and the lattice constant a. To see how this is done, see refs. 45, 48 and 6. This relation reads

$$a = \Lambda_0^{-1}(g^2\gamma_0)^{\gamma_1/(2\gamma_0^2)}\exp[-1/(2\gamma_0 g^2)][1+\mathcal{O}(g^2)]. \tag{15.186}$$

This implies that g decreases with decreasing a, in other words, for small distances the coupling constant becomes small. From (15.183) we then see that the potential is screened to zero at small distances. This is just the opposite of ordinary screening, where the potential decays rapidly for large distances. Therefore, the name 'anti-screening' has been used for this phenomenon, which is in fact the asymptotic freedom property of quarks. The constants γ_0 and γ_1 are given by $\gamma_0 = (11-2n_f/3)/(16\pi^2)$ and $\gamma_1 = (102-22n_f/3)/(16\pi^2)^2$ repectively (n_f is the number of flavours), and Λ_0 is an integration constant in this derivation – it must be fixed by experiment. Any mass is given in units of a^{-1}, which in turn is related to g through the mass constant Λ_0. The important result is that if we do not include quark masses into the theory, only a single number must be determined from experiment, and this number sets the scale for all the masses, such as the masses of glueballs, or those of massive states composed of zero-mass quarks. Therefore, after having determined Λ_0 from comparison with a single mass, all other masses and coupling constants can be determined from the theory, that is, from the simulation.

Nice this result may be, it tells us that if we simulate QCD on a lattice, and if we want the lattice constant a to be small enough to describe the continuum limit properly, we need a large lattice. The reason is that the phase diagram for the SU(3) lattice theory is simpler than that of compact QED in four dimensions. In the latter case, we have seen that there exist a Coulomb phase and a confined phase, separated by a phase transition. In lattice QCD there is only one phase, but a secret length scale is set by the lattice parameter for which (15.186) begins to hold. The lattice theory will approach the continuum theory if this equation holds, that is, if the lattice constant is sufficiently small. If we want to include a hadron in the lattice, we need a certain physical dimension to be represented by the lattice (at least a 'hadron diameter'). The small values allowed for the lattice constant and the fixed size required by the physical problem we want to describe cause the lattice to contain a very large number of sites. Whether it is allowed to take the lattice constant larger than the range where (15.186) applies is an open question, but this cannot be relied upon.

In addition to the requirement that the lattice size exceeds the hadronic scale, it must be large enough to accomodate small quark masses. The reason is that there exist excitations ('Goldstone bosons') on the scale of the quark mass. The quark masses which can currently be included are still too high too predict the instability of the ρ-meson for example.

At the time of writing, many interesting results on lattice QCD have been obtained and much is still to be expected. A very important breakthrough is the formulation of improved staggered fermion (ISF) actions, which approximate the continuum action to higher order in the latice constant than the straightforward lattice formulations discussed so far.[60–62] This makes it possible to obtain results for heavy quark, and even for lighter ones, important properties have been or are calculated,[62] such as decay constants for excited hadron states.

An interesting state of matter is the *quark-gluon plasma*, which is the QCD analog of the Kosterlitz-Thouless phase transition: the hadrons can be viewed as bound pairs or triplets of quarks, but for high densities en high temperatures, the 'dielectric' system may 'melt' into a 'conducting', dense system of quarks and gluons. This seems to have been observed very recently after some ambiguous indications. It turns out that this state of matter resembles a liquid. Lattice gauge theorists try to match these results in their large scale QCD calculations. For a recent review, see Ref. 63.

To conclude, we describe how to update gauge fields in a simulation. In a Metropolis approach we want to change the matrices $U_\mu(n)$ and then accept or reject these changes. A way to do this is to fill a list with 'random SU(3)' matrices, which are concentrated near the unit matrix. We multiply our link matrix $U_\mu(n)$ by

a matrix taken randomly from the list. For this step to be reversible, the list must contain the inverse of each of its elements. The list must be biased towards the unit matrix because otherwise the changes in the matrices become too important and the acceptance rate becomes too small. Creutz[6, 64] has developed a clever heat bath algorithm for SU(2). Cabibbo and Marinari[65] have devised an SU(3) variant of this method in which the heat bath is successively applied to SU(2) subgroups of SU(3).

Exercises

15.1 Consider the Gaussian integral

$$I_1 = \int_{-\infty}^{\infty} dx_1 \dots dx_N \, e^{-\mathbf{x}A\mathbf{x}}$$

where $\mathbf{x} = (x_1, \dots, x_N)$ is a real vector and A is a Hermitian and positive $N \times N$ matrix (positive means that all the eigenvalues λ_i of A are positive).

(a) By diagonalising A, show that the integral is equal to

$$I_1 = \sqrt{\frac{(2\pi)^N}{\prod_{i=1}^N \lambda_i}} = \sqrt{\frac{(2\pi)^N}{\det(A)}}.$$

(b) Now consider the integral

$$I_2 = \int dx_1 dx_1^* \dots dx_N dx_N^* \, e^{-\mathbf{x}^\dagger A\mathbf{x}}$$

where $\mathbf{x}$ is now a complex vector. Show that

$$I_2 = \frac{(2\pi)^N}{\det(A)}.$$

15.2 In this problem and the next we take a closer look at the free field theory. Consider the one-dimensional, periodic chain of particles with harmonic coupling between nearest neighbours, and moving in a harmonic potential with coupling constant m^2. The Lagrangian is given by

$$\mathscr{L} = \frac{1}{2} \sum_{n=-\infty}^{\infty} \left[\dot{\phi}_n^2 - (\phi_n - \phi_{n+1})^2 - m^2 \phi_n^2 \right].$$

We want to find the Hamiltonian $\mathscr{H}$ such that

$$\int [\mathscr{D}\phi_n] \, e^{-S} = \left\langle \Phi_\mathrm{i} \left| e^{-(t_\mathrm{f} - t_\mathrm{i})\mathscr{H}} \right| \Phi_\mathrm{f} \right\rangle$$

where

$$S = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathscr{L}[\phi_n(t)]\,dt$$

and the path integral $\int[\mathscr{D}\phi_n]$ is over all field configurations $\{\phi_n\}$ compatible with Φ_i at t_i and Φ_f at t_f.

We use the Fourier transforms

$$\phi_k = \sum_n \phi_n e^{ikn}; \qquad \phi_n = \int_0^{2\pi} \frac{dk}{2\pi} \phi_k e^{-ikn}.$$

(a) Show that from the fact that ϕ_n is real, it follows that $\phi_k = \phi^*_{-k}$, and that the Lagrangian can be written as

$$\mathscr{L} = \frac{1}{2}\int_0^{2\pi} \frac{dk}{2\pi} \left\{|\dot{\phi}_k|^2 - \phi_{-k}\left[m^2 + 2(1-\cos k)\right]\phi_k\right\}.$$

This can be viewed as a set of uncoupled harmonic oscillators with coupling constant $\omega_k^2 = m^2 + 2(1-\cos k)$.

(b) In Section 12.4 we have evaluated the Hamiltonian for a harmonic oscillator. Use the result obtained there to find

$$\mathscr{H} = \frac{1}{2}\int_0^{2\pi} \frac{dk}{2\pi} \left\{\hat{\pi}(k)\hat{\pi}(-k) + \hat{\phi}(-k)\left[m^2 + 2(1-\cos k)\right]\hat{\phi}(k)\right\},$$

where the hats denote operators; $\hat{\pi}(k)$ is the momentum operator conjugate to $\hat{\phi}(k)$ – they satisfy the commutation relation

$$[\hat{\pi}(k), \hat{\phi}(-k')] = i\sum_n e^{ik(k-k')n} = 2\pi\delta(k-k'),$$

where the argument of the delta-function should be taken modulo 2π.

(c) To diagonalise the Hamiltonian we introduce the operators

$$\hat{a}_k = \frac{1}{\sqrt{4\pi\omega_k}}[\omega_k\hat{\phi}(k) + i\hat{\pi}(k)];$$

$$\hat{a}_k^\dagger = \frac{1}{\sqrt{4\pi\omega_k}}[\omega_k\hat{\phi}^\dagger(k) - i\hat{\pi}^\dagger(k)].$$

Show that

$$[a_k, a_{k'}] = [a_k, a^\dagger_{-k'}] = \delta(k-k').$$

(d) Show that $\mathscr{H}$ can be written in the form

$$\mathscr{H} = \frac{1}{2}\int_0^{2\pi} dk\,\omega_k(a_k^\dagger a_k + a_k a_k^\dagger) = \int_0^{2\pi} dk\,\omega_k\left(a_k^\dagger a_k + \frac{1}{2}\right).$$

15.3 Consider the path integral for the harmonic chain of the previous problem. We have seen that the Lagrangian could be written as a k-integral over uncoupled harmonic-oscillator Lagrangians:

$$\mathscr{L} = \int_0^{2\pi} dk \mathscr{L}(k) = \frac{1}{2}\int_0^{2\pi} dk \left[|\dot{\phi}(k)|^2 - \omega_k^2 |\phi(k)|^2\right].$$

We discretise the time with time step 1 so that

$$\dot{\phi}(k,t) \to \phi(k,t+1) - \phi(k,t).$$

(a) Show that the Lagrangian can now be written as a two-dimensional Fourier integral of the form:

$$\mathscr{L} = -\frac{1}{2}\int \frac{d^2q}{(2\pi)^2}\tilde{\omega}_q^2 |\phi(q)|^2$$

with

$$\tilde{\omega}_q^2 = m^2 + 2(1-\cos q_0) + 2(1-\cos q_1);$$

q_0 corresponds to the time component and q_1 to the space-component.

(b) Show that in the continuum limit (small q), the two-point Green's function in q-space reads

$$\langle \phi_q \phi_{q'} \rangle = \frac{1}{m^2+q^2}\delta_{q,-q'}.$$

15.4 [**C**] The multigrid Monte Carlo program for the ϕ^4 field theory can be extended straightforwardly to the XY model. It is necessary to work out the coarsening of the Hamiltonian. The Hamiltonian of the XY model reads

$$\mathscr{H} = -\sum_{\langle n,n'\rangle} J\cos(\phi_n - \phi_{n'}).$$

In the coarsening procedure, the new coupling constant will vary from bond to bond, and apart from the cosines, sine interactions will be generated. The general form which must be considered is therefore

$$\mathscr{H} = -\sum_{\langle nn'\rangle} \left[J_{nn'}\cos(\phi_n - \phi_{n'}) + K_{nn'}\sin(\phi_n - \phi_{n'})\right].$$

The relation between the coarse coupling constants $J_{NN'}, K_{NN'}$ and the fine ones is

$$J_{NN'} = \sum_{nn'|NN'} \left[J_{nn'}\cos(\phi_n - \phi_{n'}) + K_{nn'}\sin(\phi_n - \phi_{n'})\right];$$

$$K_{NN'} = \sum_{nn'|NN'} \left[K_{nn'}\cos(\phi_n - \phi_{n'}) - J_{nn'}\sin(\phi_n - \phi_{n'})\right];$$

see Figure 15.5.

(a) Verify this.

(b) [**C**] Write a multigrid Monte Carlo program for the XY model. Calculate the helicity modulus using (15.97) and and check the results by comparison with Figure 15.4.

15.5 In this problem we verify that the SOR method for the free field theory satisfies detailed balance.

(a) Consider a site n, chosen at random in the SOR method. The probability distribution according to which we select a new value for the field ϕ_n in the heat bath method is

$$\rho(\phi_n) = e^{-a(\phi_n - \bar{\phi}_n)^2/2},$$

where $\bar{\phi}_n$ is the average value of the field at the neighbouring sites. In the SOR method we choose for the new value ϕ'_n at site n:

$$\phi'_n = \tilde{\phi}_n + r\sqrt{\omega(2-\omega)/a},$$

where

$$\tilde{\phi}_n = \omega\bar{\phi}_n + (1-\omega)\phi_n$$

and where r is a Gaussian random number with standard deviation 1. Show that this algorithm corresponds to a transition probability

$$T(\phi_n \to \phi'_n) \sim \exp\left[-\frac{a}{\omega(2-\omega)}\left(\phi'_n - \tilde{\phi}_n\right)^2\right].$$

(b) Show that this transition probability satisfies the detailed balance condition:

$$\frac{T(\phi_n \to \phi'_n)}{T(\phi'_n \to \phi_n)} = \frac{\exp\left[-a(\phi'_n - \bar{\phi}_n)^2/2\right]}{\exp\left[-a(\phi_n - \bar{\phi}_n)^2/2\right]}.$$

15.6 The Wilson loop correlation function for compact QED in (1+1) dimensions can be solved exactly. Links in the time direction have index $\mu = 0$, and the spatial links have $\mu = 1$. We must fix the gauge in order to keep the integrals finite. The so-called *temporal gauge* turns out convenient: in this gauge, the angles θ_0 living on the time-like bonds are zero, so that the partition sum is a sum over angles θ_1 on spatial links only. Therefore there is only a contribution from the two space-like sides of the rectangular Wilson loop. The Wilson loop correlation function is defined as

$$W(C) = \frac{\int_0^{2\pi} \prod_{n,\mu} d\theta_\mu(n) e^{\beta\cos[\sum_{n;\mu\nu}\theta_{\mu\nu}(n)]} e^{i\sum_{(n,\mu)\in C}\theta_\mu(n)}}{\int_0^{2\pi} \prod_{n,\mu} d\theta_\mu(n) e^{\beta\cos[\sum_{n;\mu\nu}\theta_{\mu\nu}(n)]}}.$$

A plaquette sum over the θ angles for a plaquette with lower-left corner at n reduces in the temporal gauge to:

$$\sum_{n;\mu\nu} \theta_{\mu\nu}(n) = \theta_1(n_0, n_1) - \theta_1(n_0 + 1, n_1).$$

(a) Show that in the temporal gauge the Wilson loop sum can be written as

$$\sum_{(n,\mu)\epsilon C} \theta_\mu(n) = \sum_{(n;\mu\nu)\epsilon A} \sum_{n;\mu\nu} \theta_{\mu\nu}(n)$$

where A is the area covered by the plaquettes enclosed by the Wilson loop.

(b) Use this to show that the Wilson loop correlation function factorises into a product of plaquette-terms. Defining

$$\theta_P(n) = \sum_{n;\mu\nu} \theta_{\mu\nu}(n),$$

where P denotes the plaquettes, we can write:

$$W(C) = \frac{\int \prod_P d\theta_P \exp\left[\beta \cos\theta_P + i\theta_P\right]}{\int \prod_P d\theta_P \exp\left[\beta \cos\theta_P\right]}.$$

(c) Show that this leads to the final result:

$$W(C) = \left[\frac{I_1(\beta)}{I_0(\beta)}\right]^A$$

where $I_n(x)$ is the modified Bessel function and A is the area enclosed by the Wilson loop.

References

[1] R. Balian and J. Zinn-Justin, eds., *Méthodes en théorie des champs / Methods in Field theory*, vol. XXVIII of *Les Houches Summer School Proceedings*. Amsterdam, North-Holland, 1975.

[2] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*. New York, McGraw-Hill, 1980.

[3] S. Weinberg, *The Quantum Theory of Fields*, vol. 1 and 2. Cambridge, Cambridge University Press, 1995.

[4] D. Bailin and A. Love, *Introduction to Gauge Field Theory*. Bristol, Adam Hilger, 1986.

[5] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*. New York, Oxford University Press, third ed., 1996.

[6] M. Creutz, *Quarks, Gluons and Lattices*. Cambridge, Cambridge University Press, 1983.

[7] M. Lüscher and P. Weisz, 'Scaling laws and triviality bounds in the lattice-ϕ^4 theory. 1. One-component model in the symmetric phase,' *Nucl. Phys. B*, **290,** 25–60, 1987.

[8] M. Lüscher, 'Volume dependence of the energy spectrum in massive quantum field theories (I). Stable particle states,' *Comm. Math. Phys.*, **104,** 177–206, 1986.

[9] J. Goodman and A. D. Sokal, 'Multigrid Monte Carlo method for lattice field theories,' *Phys. Rev. Lett.*, **56,** 1015–1018, 1986.

[10] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*. Cambridge, Cambridge University Press, second ed., 1992.

[11] S. L. Adler, 'Over-relaxation method for the Monte Carlo evaluation of the partition function for multiquadratic actions,' *Phys. Rev. Lett.*, **23,** 2901–2904, 1981.

[12] C. Whitmer, 'Over-relaxation methods for Monte Carlo simulations of quadratic and multiquadratic actions,' *Phys. Rev. D*, **29,** 306–311, 1984.

[13] S. Duane and J. B. Kogut, 'Hybrid stochastic differential-equations applied to quantum chromodynamics,' *Phys. Rev. Lett.*, **55,** 2774–2777, 1985.

[14] S. Duane, 'Stochastic quantization versus the microcanonical ensemble – getting the best of both worlds,' *Nucl. Phys. B*, **275,** 398–420, 1985.

[15] E. Dagotto and J. B. Kogut, 'Numerical analysis of accelerated stochastic algorithms near a critical temperature,' *Phys. Rev. Lett.*, **58,** 299–302, 1987.

[16] E. Dagotto and J. B. Kogut, 'Testing accelerated stochastic algorithms in 2 dimensions - the SU(3)×SU(3) spin model,' *Nucl. Phys. B*, **290,** 415–468, 1987.

[17] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, 'Langevin simulation of lattice field theories,' *Phys. Rev. D*, **32,** 2736–2747, 1985.

[18] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, 'Hybrid Monte Carlo,' *Phys. Lett. B*, **195,** 216–222, 1987.

[19] M. Creutz, 'Global Monte Carlo algorithms for many-fermion systems,' *Phys. Rev. D*, **38,** 1228–1238, 1988.

[20] R. H. Swendsen and J.-S. Wang, 'Nonuniversal critical dynamics in Monte Carlo simulations,' *Phys. Rev. Lett.*, **58,** 86–88, 1987.

[21] U. Wolff, 'Comparison between cluster Monte Carlo algorithms in the Ising models,' *Phys. Lett. B*, **228,** 379–382, 1989.

[22] C. M. Fortuin and P. W. Kasteleyn, 'On the random cluster model. I. Introduction and relation to other models,' *Physica*, **57,** 536–564, 1972.

[23] J. Hoshen and R. Kopelman, 'Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm,' *Phys. Rev. B*, **14,** 3438–3445, 1976.

[24] U. Wolff, 'Monte Carlo simulation of a lattice field theory as correlated percolation,' *Nucl. Phys. B*, **FS300,** 501–516, 1988.

[25] M. Sweeny, 'Monte Carlo study of weighted percolation clusters relevant to the Potts models,' *Phys. Rev. B*, **27,** 4445–4455, 1983.

[26] M. Creutz, 'Microcanonical Monte Carlo simulation,' *Phys. Rev. Lett.*, **50,** 1411–1414, 1983.

[27] U. Wolff, 'Collective Monte Carlo updating for spin systems,' *Phys. Rev. Lett.*, **69,** 361–364, 1989.

[28] R. G. Edwards and A. D. Sokal, 'Dynamic critical behaviour of Wolff's collective-mode Monte Carlo algorithm for the two-dimensional O(n) nonlinear σ model,' *Phys. Rev. D*, **40,** 1374–1377, 1989.

[29] B. Nienhuis, 'Coulomb gas formulation of two-dimensional phase transitions,' vol. 11 of *Phase Transitions and Critical Phenomena*, London, Academic Press, 1987.

[30] D. J. Bishop and J. D. Reppy, 'Study of the superfluid transition in two-dimensional ^{4}He films,' *Phys. Rev. Lett.*, **40,** 1727–1730, 1978.

[31] C. J. Lobb, 'Phase transitions in arrays of Josephson junctions,' *Physica B*, **126,** 319–325, 1984.

[32] T. Ohta and D. Jasnow, 'XY model and the superfluid density in two dimensions,' *Phys. Rev. B*, **20,** 139–146, 1979.

[33] M. Kosterlitz and D. J. Thouless, 'Ordering, metastability and phase transitions in two-dimensional systems,' *J. Phys. C*, **6,** 1181–1203, 1973.

[34] M. Plischke and H. Bergersen, *Equilibrium Statistical Physics*. Englewood Cliffs, NJ, Prentice-Hall, 1989.

[35] C. Dress and W. Krauth, 'Cluster Algorithm for hard spheres and related systems,' *J. Phys. A*, **28,** L597–601, 1995.

[36] J. Liu and E. Luijten, 'Rejection-free geometric cluster algorithm for complex fluids,' *Phys. Rev. Lett.*, **92,** 035504 (4 pages), 2004.

[37] J. R. Heringa and H. W. J. Blöte, 'Geometric cluster Monte Carlo simulation,' *Phys. Rev. E*, **57,** 4976–4978, 1998.

[38] J. Goodman and A. D. Sokal, 'Multigrid Monte Carlo method. Conceptual foundations,' *Phys. Rev. D*, **40,** 2035âĂŞ2071, 1989.

[39] R. G. Edwards, J. Goodman, and A. D. Sokal, 'Multigrid Monte Carlo (II). Two-dimensional XY model,' *Nucl. Phys. B*, **354,** 289–327, 1991.

[40] J. D. Jackson, *Classical Electrodynamics*. New York, John Wiley, second ed., 1974.

[41] J. W. Negele and H. Orland, *Quantum Many-particle Systems*. Frontiers in Physics, Redwood City, Addison-Wesley, 1988.

[42] F. Wegner,, 'Duality in generalized Ising models and phase transitions without local order parameters,' *J. Math. Phys.*, **12,** 2259–2272, 1971.

[43] K. G. Wilson, 'Confinement of quarks,' *Phys. Rev. D*, **10,** 2445–2459, 1974.

[44] J. B. Kogut, 'An introduction to lattice gauge theory and spin systems,' *Rev. Mod. Phys.*, **51,** 659–713, 1979.

[45] H. J. Rothe, *Lattice Gauge Theories – An Introduction*, vol. 43 of *World Scientific Lecture Notes in Physics*. Singapore, World Scientific, 1992.

[46] A. H. Guth, 'Existence proof of a nonconfining phase in four-dimensional U(1)

lattice gauge theory,' *Phys. Rev. D*, **21,** 2291–2307, 1980.

[47] G. 't Hooft, 'Unpublished remarks at the Marseille conference on gauge theories,' 1972.

[48] J. B. Kogut, 'The lattice gauge theory approach to quantum chromodynamics,' *Rev. Mod. Phys.*, **55,** 775–836, 1983.

[49] H. B. Nielsen and M. Ninomiya, 'Absence of neutrinos on a lattice. 1. Proof by homotopy theory,' *Nucl. Phys. B*, **185,** 20–40, 1981.

[50] J. B. Kogut and L. Susskind, 'Hamiltonian form of Wilson's lattice gauge theories,' *Phys. Rev. D*, **11,** 395–408, 1975.

[51] T. J. Banks, R. Myerson, and J. B. Kogut,, 'Phase transitions in abelian lattice gauge theories,' *Nucl. Phys. B*, **129,** 493–510, 1977.

[52] J. Kuti, 'Lattice field theories and dynamical fermions,' in *Computational Physics. Proceedings of the Thirty-second Scottish University Summer School in Physics.* (R. D. Kenway and G. S. Pawley, eds.), Nato ASI, pp. 311–378, 1987.

[53] F. Fucito and G. Marinari, 'A stochastic approach to simulations of fermionic systems,' *Nucl. Phys. B*, **190,** 266–278, 1981.

[54] G. Bahnot, U. M. Heller, and I.O. Stamatescu, 'A new method for fermion Monte Carlo,' *Phys. Lett. B*, **129,** 440–444, 1983.

[55] J. Polonyi and H. W. Wyld, 'Microcanonical simulation of fermion systems,' *Phys. Rev. Lett.*, **51,** 2257–2260, 1983.

[56] O. Martin and S. Otto, 'Reducing the number of flavors in the microcanonical method,' *Phys. Rev. D*, **31,** 435–437, 1985.

[57] S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, 'Hybrid molecular dynamics algorithm for the numerical simulation of quantum chromodynamics,' *Phys. Rev. D*, **35,** 2531–2542, 1988.

[58] D. J. Gross and F. Wilczek, 'Ultra-violet behavior of non-abelian gauge theories,' *Phys. Rev. Lett.*, **30,** 1343–1346, 1973.

[59] H. D. Politzer, 'Reliable perturbation results for strong interactions,' *Phys. Rev. Lett.*, **30,** 1346–1349, 1973.

[60] S. Naik, 'On-shell improved action for QCD with Susskind fermions and the asymptotic freedom scale,' *Nucl. Phys. B*, **316,** 238–268, 1989.

[61] G. P. Lepage, 'Flavor-symmetry restoration and Szymanzik improvement for staggered quarks,' *Phys. Rev. D*, **59,** 074502 (4 pages), 1999.

[62] C. T. H. Davies and G. P.Lepage, 'Lattice QCD meets experiment in hadron physics.' hep-lat/0311041, 2003.

[63] O. Philipsen, 'The QCD phase diagram at zero and small baryon density.' hep-lat/0510077, 2005.

[64] M. Creutz, 'Monte Carlo study of quantized SU(2) gauge theory,' *Phys. Rev. D*, **21,** 2308–2315, 1980.

[65] N. Cabibbo and E. Marinari, 'A new method for updating $SU(N)$ matrices in computer-simulations of gauge-theories,' *Phys. Lett. B*, **119,** 387–390, 1982.

ß