Lecture 16: LATTICE GAUGE THEORIES: AN INTRODUCTION

In this chapter, we discuss gauge theories on the lattice (chosen hypercubic for convenience), which can be considered as non-perturbative regularizations of the continuum gauge theories studied in chapters 13–15: the low temperature or small coupling expansion of the lattice model is a regularized continuum perturbation theory.

We concentrate on pure lattice gauge theories (without fermions). Physically, this means that we cannot investigate many properties of a realistic theory like QCD where fermions are coupled through a gauged $SU(3)$ colour group, but we can still investigate with lattice methods one important question:

Does the theory generate confinement, that is, a force between charged particles increasing at large distances, so that heavy quarks in the fundamental representation cannot be separated?

More generally, can one find charged (from the gauge group point of view) asymptotic states like massless vector particles in the theory?

Other problems which we do not consider here, can also be discussed in this framework: for example, the appearance of massive group singlet bound states in the spectrum (gluonium), the question of a deconfinement transition at finite physical temperature in QCD.

We first construct lattice models with gauge symmetry and show that gauge fields are replaced by parallel transporter. We study pure gauge theories (without matter fields) on the lattice. We discover that gauge theories have properties quite different from the ferromagnetic systems we have studied so far. In particular, the absence of a local order parameter will force us to examine the behaviour of a non-local quantity, a functional of loops called hereafter Wilson's loop to distinguish between the confined and deconfined phases. Results will be obtained in the high temperature or strong coupling limit and in the mean field approximation. However, we emphasize that the main physics results of lattice gauge theories are obtained, after dynamical quarks are added, from large scale numerical

simulations with realistic values of physical parameters (including small quark masses).

16.1 Gauge invariance on the lattice

The construction of lattice gauge theories is based on the idea of parallel transport that has already been introduced in chapter 14.

We start from a model possessing a global (rigid) symmetry group G, and we want to make it gauge invariant.

To each site i of a lattice, we associate a set of dynamic variables $\boldsymbol{\varphi}_i$, representing matter fields, on which acts an orthogonal representation $\mathcal{D}(G)$ of the group G:

$$\boldsymbol{\varphi}_{\mathbf{g}} = \mathbf{g}\boldsymbol{\varphi}\,, \qquad \mathbf{g} \in \mathcal{D}(G)\,.$$

A model is gauge invariant (local invariance) if it is invariant under independent group transformation on each lattice site i. For the φ-measure of

integration as well as for all the terms in the lattice action which depend only on one site, global invariance implies local invariance as in the continuum. Problems arise only with terms that connect different lattice sites.

Let us consider, for example, a term in the action of the form $\boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_j$, i and j being different sites on the lattice. Such a term is invariant under global but not local transformations:

$$\boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_j \mapsto \boldsymbol{\varphi}_i \mathbf{g}_i^T \, \mathbf{g}_j \boldsymbol{\varphi}_j \,,$$

(where $\mathbf{g}^T$ means $\mathbf{g}$ transposed). To render it invariant, it is necessary to introduce a new dynamic variable, a matrix $\mathbf{U}_{ij}$ belonging to the representation $\mathcal{D}(G)$, which depends on the two sites i, j and transforms like

$$\mathbf{U}_{ij} \mapsto \mathbf{g}_i \mathbf{U}_{ij} \mathbf{g}_j^T \,. \tag{16.1}$$

Then, the quantity

$$\boldsymbol{\varphi}_i \mathbf{U}_{ij} \boldsymbol{\varphi}_j \tag{16.2}$$

is gauge invariant. Moreover, if $\mathbf{U}_{ij}$ and $\mathbf{U}_{jk}$ are two matrices transforming with the rule (16.2) then the product of matrices $\mathbf{U}_{ij}\mathbf{U}_{jk}$ transforms like

$$\mathbf{U}_{ij}\mathbf{U}_{jk} \mapsto \mathbf{g}_i \mathbf{U}_{ij}\mathbf{U}_{jk}\mathbf{g}_k^T . \tag{16.3}$$

In the transformation (16.1), we recognize the transformation of a parallel transporter. In the continuum, a parallel transporter depends not only on the end-points i, j but also on the curve joining them. Moreover, in a local field theory one needs only transport along infinitesimal curves which can be expressed in terms of a gauge field or connection, element of the representation of the Lie algebra.

On the lattice curves follow links, the segments which connect adjacent sites. The minimum displacement is a link and two arbitrary lattice sites can be joined by a path formed of links of the lattice. As a consequence of the composition rule (16.3), one can thus take as dynamic variables elements $\mathbf{U}_\ell$ of the group representation associated with parallel transport along oriented

links of the lattice, which transform like

$$\mathbf{U}_\ell \equiv \mathbf{U}_{ab} \mapsto \mathbf{g}_a \mathbf{U}_\ell \, \mathbf{g}_b^T \,,$$

where the link ℓ goes from site b to a. It is consistent with the transformation law to choose

$$\mathbf{U}_{ba} = {}^T\mathbf{U}_{ab} = \mathbf{U}_{ab}^{-1}. \tag{16.4}$$

Then, we can choose for matrix $\mathbf{U}_{ij}$ any parallel transporter product of link variables along a path C_{ij} joining j to i:

$$\mathbf{U}[C_{ij})] = \prod_{\text{links } \ell \in C_{ij}} \mathbf{U}_\ell \,,$$

where the product is ordered along the path.

Relation with the continuum formulation: the Abelian example. In continuum field theory, in the Abelian $U(1)$ example, we have already explicitly constructed the parallel transporter (equation (13.25)) which is an element of the $U(1)$ group. In terms of the gauge field A_μ, it reads

$$U[C_{xy}] = \exp\left[-ie \oint_C \sum_\mu A_\mu(s) \mathrm{d}s_\mu\right],$$

in which e is the gauge coupling constant, and the gauge field is integrated over an oriented piecewise differentiable curve going from x to y. Indeed, in a gauge transformation a charged field φ and the gauge field

$$\varphi(x) \mapsto \mathrm{e}^{i\Lambda(x)}\, \varphi(x), \quad A_\mu(x) \mapsto A_\mu(x) - \frac{1}{e}\partial_\mu \Lambda(x)$$

and, thus,

$$e \int_C \sum_\mu A_\mu(s) \mathrm{d}s_\mu \mapsto e \int_C \sum_\mu A_\mu(s) \mathrm{d}s_\mu - \Lambda(y) + \Lambda(x),$$

the transformation of $U[C_{xy}]$ is

$$U[C_{xy}] \mapsto \mathrm{e}^{i\Lambda(y)-i\Lambda(x)}\, U[C_{xy}].$$

The non-Abelian case. In the non-Abelian case, the explicit relation is more complicated because the gauge field $\sum_\alpha A_\mu^\alpha(x)\mathbf{t}^\alpha$ is an element of the Lie algebra of G and the matrices representing the field at different points do not commute. It can be formally written as

$$\mathbf{U}[C_{xy}] = \mathrm{P}\left\{\exp\left[\oint_C \sum_{\alpha,\mu} A_\mu^\alpha(s)\mathbf{t}^\alpha \mathrm{d}s_\mu\right]\right\},$$

in which the symbol P means path-ordered integral.

16.2 The pure gauge theory

We now discuss the pure gauge theory and its formal continuum limit as obtained from a low temperature, strong coupling expansion.

16.2.1 Action and partition function

We now have to construct a gauge invariant interaction for the gauge elements $\mathbf{U}$. It follows from the transformation (16.1) that only the traces of the products of $\mathbf{U}$'s on closed loops are gauge invariant. On a hypercubic lattice, the shortest loop is a square, called hereafter a plaquette. In what follows, we thus consider a pure gauge action of the form

$$\mathcal{S}(\mathbf{U}) = -\sum_{\text{plaquettes}} \operatorname{tr} \mathbf{U}_{ij}\mathbf{U}_{jk}\mathbf{U}_{kl}\mathbf{U}_{li}\,. \tag{16.5}$$

The appearance of products of parallel transporters along closed loops is not surprising since we know quite generally that the curvature tensor $\mathbf{F}_{\mu\nu}$

which appears in the pure gauge action of the continuum theory is associated with infinitesimal transport along a closed loop. Note that each plaquette appears with both orientations in such a way that the sum is real when the group is unitary.

The partition function. We can then write a partition function corresponding to the action (16.5) as

$$\mathcal{Z} = \int \prod_{\text{links}\{ij\}} \mathrm{d}\mathbf{U}_{ij} \; \mathrm{e}^{-\beta_p \mathcal{S}(\mathbf{U})}, \tag{16.6}$$

in which β_p is the plaquette coupling. We integrate over $\mathbf{U}_{ij}$ with the group invariant (de Haar) measure associated with the group G. In contrast to continuum gauge theories, the expression (16.6) is well-defined on the lattice (at least as long as the volume is finite) because the group is compact and thus the volume of the group is finite. Therefore, gauge fixing is not required and a completely gauge invariant formulation of the theory is possible.

16.2.2 Low temperature analysis

To understand the precise connection between the lattice theory (16.6) and the continuum field theory, we investigate the lattice theory at low temperature, that is, at large positive β_p. In this limit, the partition function is dominated by minimal energy configurations.

Let us show that the minimum of the energy corresponds to matrices $\mathbf{U}$ gauge transform of the identity. We start from a first plaquette 1234. Without loss of generality, we can set

$$\mathbf{U}_{12} = \mathbf{g}_1^{-1}\mathbf{g}_2\,, \qquad \mathbf{g}_1,\, \mathbf{g}_2 \in \mathcal{D}(G).$$

The matrix $\mathbf{g}_1$ is arbitrary and $\mathbf{g}_2$ is calculated from $\mathbf{U}_{12}$ and $\mathbf{g}_1$. Then, we can also set

$$\mathbf{U}_{23} = \mathbf{g}_2^{-1}\mathbf{g}_3\,, \quad \mathbf{U}_{34} = \mathbf{g}_3^{-1}\mathbf{g}_4\,.$$

These relations define first $\mathbf{g}_3$, then $\mathbf{g}_4$. The minimum of the action is obtained when the real part of all traces is maximum, that is, when the prod-

ucts of the group elements on a plaquette are 1. (The trace of an orthogonal matrix $\mathbf{U}$ is maximum when all its eigenvalues are 1.) In particular,

$$\mathbf{U}_{12}\mathbf{U}_{23}\mathbf{U}_{34}\mathbf{U}_{41} = \mathbf{1}\,,$$

which yields

$$\mathbf{U}_{41} = \mathbf{g}_4^{-1}\mathbf{g}_1\,.$$

If we now take an adjacent plaquette the argument can be repeated for all links but one, which has already been fixed. In this way, we can show that the minimum of the action is a pure gauge. Thus, when the coupling constant β_p becomes very large, all group elements are constrained to stay, up to a gauge transformation, close to the identity (in a finite volume with consistent boundary conditions). From this analysis, we learn that the minimum of the potential is highly degenerate at low temperature, since it is parametrized by a gauge transformation, which corresponds to a finite

number of degrees of freedom per site. This unusual property of lattice gauge theories corresponds to the property that the gauge action in classical mechanics determines the motion only up to a gauge transformation. To perform a low temperature expansion, it becomes necessary to 'fix' the gauge in order to sum over all minima.

Low temperature expansion. We choose a gauge such that the minimum of the energy corresponds to all matrices $\mathbf{U} = \mathbf{1}$. At low temperature, the matrices $\mathbf{U}$ are then close to the identity:

$$\mathbf{U}(x, x + an_\mu) = 1 - a\mathbf{A}_\mu(x) + O\left(a^2\right),$$

in which a is the lattice spacing, x the point on the lattice, and n_μ the unit vector in the direction μ. We know from the discussion of section 14.1 that the matrix $\mathbf{A}_\mu(x)$ is the connection or gauge field. We have already shown that the transformation (16.1) of the parallel transporter implies for $\mathbf{A}_\mu(x)$

at leading order in the lattice spacing:

$$\mathbf{A}_\mu(x) \mapsto \mathbf{g}(x)\partial_\mu \mathbf{g}^{-1}(x) + \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x),$$

which is the usual gauge transformation.

We now expand the lattice action for small fields. To simplify calculations, we parametrize the orthogonal matrix $\mathbf{U}$ associated with the link $(x, x+an_\mu)$ in terms of the antisymmetric matrix $\mathbf{A}_\mu(x)$ as

$$\ln \mathbf{U}(x + an_\mu, x) = -a\mathbf{A}_\mu(x + \tfrac{1}{2}an_\mu) + O\left(a^3\right). \tag{16.7}$$

We verify below that we need $\mathbf{U}$ up to order a^2. With the parametrization (16.7), equation (16.4) implies that the term of order a^2 vanishes. We now define the antisymmetric matrix $\mathbf{F}_{\mu\nu}(x)$ by

$$\begin{aligned} \mathrm{e}^{-a^2\mathbf{F}_{\mu\nu}(x)} = {} & \mathbf{U}(x, x + an_\nu)\mathbf{U}(x + an_\nu, x + a(n_\mu + n_\nu)) \\ & \times \mathbf{U}(x + a(n_\mu + n_\nu), x + an_\mu)\mathbf{U}(x + an_\mu, x). \end{aligned} \tag{16.8}$$

To calculate $\mathbf{F}_{\mu\nu}(x)$, we introduce the expansion (16.7) and use repeatedly the Baker–Hausdorf formula:

$$\ln\left(\mathrm{e}^{A}\,\mathrm{e}^{B}\right) = A + B + \tfrac{1}{2}\,[A,B] + \cdots\ .$$

Applied to the product of several factors, it takes the form

$$\ln\left(\mathrm{e}^{A_1}\,\mathrm{e}^{A_2}\ldots\mathrm{e}^{A_n}\right) = \sum_i A_i + \tfrac{1}{2}\sum_{i<j}[A_i,A_j] + \cdots\ ,$$

and, therefore,

$$\begin{aligned} a^2\mathbf{F}_{\mu\nu}(x) = a\,\big[&\mathbf{A}_\mu(x+\tfrac{1}{2}an_\mu) + \mathbf{A}_\nu(x+an_\mu+\tfrac{1}{2}an_\nu)\\ &-\mathbf{A}_\mu(x+an_\nu+\tfrac{1}{2}an_\mu) - \mathbf{A}_\nu(x+\tfrac{1}{2}an_\nu)\big] + a^2\,[\mathbf{A}_\mu(x),\mathbf{A}_\nu(x)] + O\left(a^3\right). \end{aligned}$$

At leading order, we recover the curvature tensor

$$\mathbf{F}_{\mu\nu}(x) = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu + [\mathbf{A}_\mu,\mathbf{A}_\nu] + O\left(a\right).$$

We obtain one term in the plaquette action by taking the trace of expression (16.8). Since $\mathbf{F}_{\mu\nu}$ is an antisymmetric matrix $\mathrm{tr}\,\mathbf{F}_{\mu\nu}$ vanishes. Thus,

$$\mathrm{tr}\,\mathrm{e}^{-a^2\mathbf{F}_{\mu\nu}(x)} = \mathrm{tr}\,\mathbf{1} + a^4\,\mathrm{tr}\,\mathbf{F}_{\mu\nu}^2(x) + O\left(a^6\right).$$

This result shows that the leading term of the small field expansion of the plaquette action (16.5) is the standard gauge action studied in chapter 14. The relation between β_p and the bare coupling constant e_0 of continuum gauge theories is thus

$$a^4\beta_p \sim e_0^{-2}\,. \tag{16.9}$$

As anticipated in chapters 13,14, we conclude that the low temperature expansion, in a fixed gauge, of lattice gauge theories provides indeed a lattice regularization of continuum gauge theories. We have here discussed only the pure gauge action, but the generalization to matter fields is simple. Higher order terms in the small field expansion yield additional interactions

needed to maintain gauge invariance on the lattice. This is not surprising: we have already shown that the gauge invariant extension of Pauli–Villars's regularization also introduces additional interactions.

16.3 Wilson's loop and confinement

In section 15.4.2, we have calculated the RG β-functions for non-Abelian gauge theories and shown that pure gauge theories are asymptotically free in four dimensions, which means that the origin in the coupling constant space is an UV fixed point and also implies that the effective interaction increases at large distance. Therefore, as in the case of the 2D non-linear σ-model, the spectrum of a non-Abelian gauge theory cannot be determined from perturbation theory. To explain the non-observation of free quarks, it has been conjectured that the spectrum of the symmetric phase consists only in neutral states, that is, states which are singlets for the group transformations.

Clearly, it would be convenient to identify a local order parameter, that is, a local observable whose expectation value would distinguish between the QED phase of Abelian gauge theories, in which charge states can be produced, from the so-called confined phase. However, in gauge theories such a local order parameter does not exist (see Elitzur's theorem). This property follows from the simple remark that physical observables correspond to gauge invariant operators which are neutral by construction. Moreover, we have seen in the study of continuum gauge theories (chapters 13, 14) that the only gauge independent quantities corresponding to non-gauge invariant operators are the S-matrix elements. Since it is very difficult to determine S-matrix elements beyond perturbation theory, it has been suggested by Wilson to study, in pure gauge theories, a gauge invariant non-local quantity, the energy of the vacuum in presence of largely separated static charges. We thus first study this quantity in pure Abelian gauge theories, in which, in the continuum, all calculations can be done explicitly.

16.3.1 Wilson's loop in continuum Abelian gauge theories

In continuum field theory, in order to calculate the average energy, it is necessary to introduce the gauge Hamiltonian, and, therefore, convenient to work in the temporal gauge. We have constructed a wave function for two static point-like charges, in the temporal gauge, in section 13.4 (equation (13.38)):

$$\psi(A) = \exp\left[-ie \oint_{C_0} \sum_i A_i(s) \mathrm{d}s_i\right],$$

in which the charges are located at both ends of the curve C_0.

By evaluating the behaviour for large time T of the matrix element

$$W(C_0) = \left\langle \psi \left| \mathrm{e}^{-HT} \right| \psi \right\rangle,$$

in which H is the gauge Hamiltonian in the temporal gauge, we obtain the energy $E(C_0)$ of the vacuum in presence of static charges:

$$W(C_0) \underset{T\to\infty}{\sim} \mathrm{e}^{-TE(C_0)}.$$

If the charges are separated by a distance R, we expect E to depend only on R and not on C_0.

The loop functional $W(C_0)$ can be calculated from a field integral:

$$W(C_0) = \left\langle \exp\left[-ie \oint_{C_0'} \sum_\mu A_\mu(s) \mathrm{d}s_\mu\right]\right\rangle,$$

C_0', which is now defined in space and time, is the union of two curves, which coincide with C_0 at time 0, and with $-C_0$ at time T, respectively. The expectation value here means average over gauge field configurations.

Since in the temporal gauge the time component of A_μ vanishes, we can add to C_0' two straight lines in the time direction which join the ends of the curves $C_0(t=0)$ and $C_0(t=T)$. $W(C_0)$ then becomes a functional of a closed loop C (see Fig. 16.1):

$$W(C_0) \equiv W(C) = \left\langle \exp\left[-ie \oint_{C} \sum_\mu A_\mu(s) \mathrm{d}s_\mu\right]\right\rangle. \tag{16.10}$$

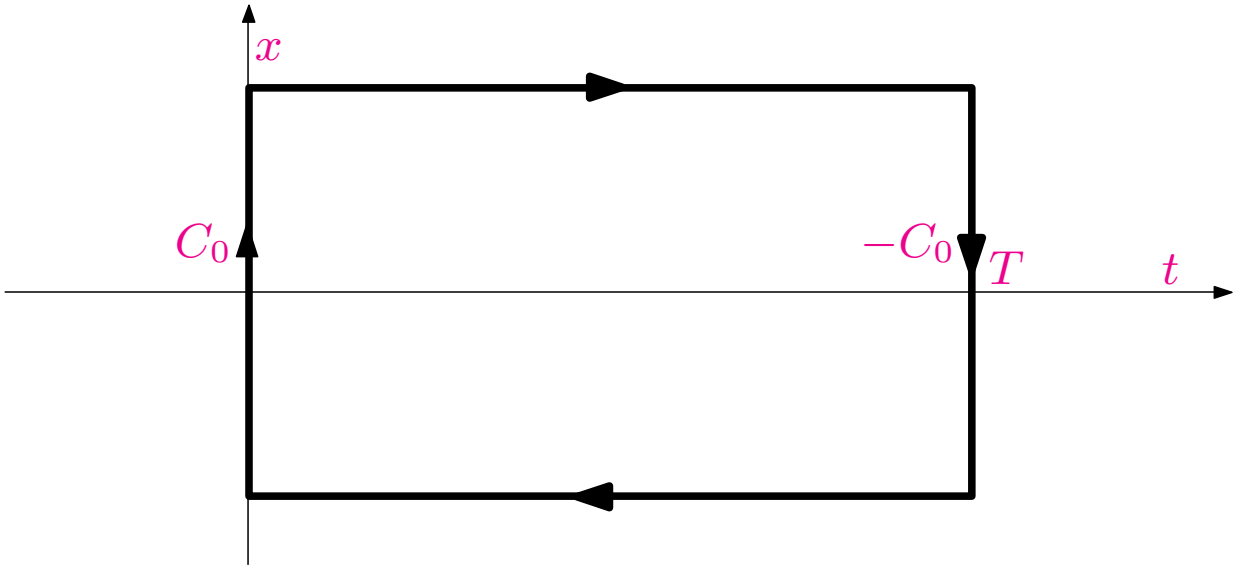

Fig. 16.1 – The loop C.

The advantage of the representation (16.10) is that it is explicitly gauge invariant since it is the expectation value of the parallel transporter corresponding to a closed loop in space–time.

The question of confinement is related to the behaviour of the energy E when the separation R between charges becomes large. In a pure Abelian gauge theory, in the continuum, which is a free field theory, the expression

(16.10) can be evaluated explicitly. To simplify calculations we take for C_0 also a straight line and use Feynman's gauge. The quantity $W(C)$ then is given by

$$W(C) = \int [\mathrm{d}A_\mu] \exp\left[-\mathcal{S}(A) + \int \mathrm{d}^d x \sum_\mu J_\mu(x) A_\mu(x)\right]$$

with

$$\mathcal{S}(A) = \tfrac{1}{2} \int \mathrm{d}^d x \sum_{\mu,\nu} [\partial_\mu A_\nu(x)]^2$$

and

$$J_\mu(x) = -ie \oint_C \delta(x - s) \mathrm{d}s_\mu \,.$$

The result is

$$\ln W(C) = -\frac{\Gamma(d/2-1)}{8\pi^{d/2}} \, \mathrm{e}^2 \oint_{C\times C} \mathrm{d}\mathbf{s}_1 \cdot \mathrm{d}\mathbf{s}_2 \, |\mathbf{s}_1 - \mathbf{s}_2|^{2-d} \,. \tag{16.11}$$

The integral in the right hand side exhibits a short distance singularity, and a short distance cut-off thus is required. Moreover, to normalize the right hand side of equation (16.11), we divide it by $W(C)$ taken for $R = a$, a being a fixed distance. We now write more explicitly the integrals:

$$\oint_{C\times C} \frac{\mathrm{d}\mathbf{s}_1 \cdot \mathrm{d}\mathbf{s}_2}{2\,|\mathbf{s}_1 - \mathbf{s}_2|^{d-2}} = \int_0^T |u-t|^{2-d}\,\mathrm{d}u\,\mathrm{d}t + \int_0^R |x-y|^{2-y}\,\mathrm{d}x\,\mathrm{d}y$$
$$- \int_0^R \left[(x-y)^2 + T^2\right]^{1-d/2}\mathrm{d}x\,\mathrm{d}y - \int_0^T \left[(t-u)^2 + R^2\right]^{1-d/2}\mathrm{d}t\,\mathrm{d}u\,.$$

The first term in the right hand side is cancelled by the normalization. The second term is independent of T and, therefore, negligible for large T. It is actually related to the scalar product of the wave function $\psi(A)$ and the ground state eigenfunction. The third term decreases with T for $d > 2$

which we now assume. Only the last term increases with T:

$$\int_0^T \left\{ \left[(t-u)^2 + R^2\right]^{1-d/2} - \left[(t-u)^2 + a^2\right]^{1-d/2} \right\} \mathrm{d}t\, \mathrm{d}u$$
$$\sim \sqrt{\pi}\, \frac{\Gamma\big((d-3)/2\big)}{\Gamma(d/2-1)} \left(R^{3-d} - a^{3-d}\right) T\,.$$

Therefore, the vacuum energy $E(R)$ in presence of the static charges has the form

$$E(R) - E(a) = \frac{e^2}{4\pi^{(d-1)/2}} \Gamma\big((d-3)/2\big) \left(a^{3-d} - R^{3-d}\right).$$

We recognize the Coulomb potential between two charges.

For $d \leq 3$, the energy of the vacuum increases without bound when the charges are separated, and free charges cannot exist.

For $d = 3$, the potential increases logarithmically.

For $d = 2$, the Coulomb potential increases linearly with distance.

In more general situations, the method that we have used above to determine the energy is complicated because we have to take the large T limit first and then evaluate the large R behaviour. It is more convenient to take a square loop, $T = R$, and evaluate the large R behaviour of $W(C)$. Here, we obtain

$$\begin{aligned}
&\ln W\left[C(R)\right] - \ln W\left[C(a)\right] \\
&\quad = \frac{1}{2\pi^{d/2}}\Gamma(d/2-1)e^2\Bigg\{\int_0^R \left[(u-t)^2 + R^2\right]^{1-d/2} \mathrm{d}u\,\mathrm{d}t \\
&\qquad - \int_0^a \left[(u-t)^2 + a^2\right]^{1-d/2} \mathrm{d}u\,\mathrm{d}t - \int_a^R |u-t|^{2-d}\mathrm{d}u\,\mathrm{d}t\Bigg\}.
\end{aligned}$$

For $d > 3$, dimensions in which the Coulomb potential decreases, the right hand side is dominated by terms which correspond to the region $|s_1 - s_2| \ll R$ in equation (16.11):

$$\ln W\left[C(R)\right] - \ln W\left[C(a)\right] \sim \text{const.} \times R\,.$$

This is called the perimeter law since $\ln W(C)$ is proportional to the perimeter of C and is, therefore, relevant to the $d = 4$ Coulomb phase.

By contrast, for $d \leq 3$, $\ln W(C)$ increases as R^{4-d}. The reason is that two charges separated on C by a distance of order R, feel a potential of order R^{d-3}.

In particular for $d = 2$, $\ln W(C)$ increases like R^2, that is, like the area of the surface enclosed by C: this is the area law expected in confinement situations.

16.3.2 Non-Abelian gauge theories

It follows from the discussion of section 14.3 that in the temporal gauge the wave function corresponding to two opposite point-like static charges is also related to a parallel transporter along a curve joining the charges. The same arguments as in the Abelian case, show that the expectation value of the operator e^{-TH} in the corresponding state is given by the average, in the

sense of the field integral, of the parallel transporter along a closed loop:

$$W(C) = \left\langle \mathrm{P} \exp\left[-i \oint_C \sum_\mu \mathbf{A}_\mu(s) \mathrm{d}s_\mu \right] \right\rangle,$$

in which we recall that the symbol P means path ordering since the matrices $\mathbf{A}_\mu(s)$ at different points do not commute.

If we calculate $W(C)$ in perturbation theory, we find of course at leading order the same results as in the Abelian case. However, we know from renormalization group, that we cannot trust perturbation theory at large distances. Therefore, to get a qualitative idea about the phase structure we first use the lattice model to calculate $W(C)$ in the large coupling or high temperature limit $\beta_p \to 0$.

Wilson's loop: strong coupling expansion. Here, we assume that the group we consider has a non-trivial centre. We take the explicit example of gauge elements on the lattice belonging to the fundamental representation

of $SU(N)$ (whose centre is $\mathbb{Z}_N$, with elements the identity multiplied by roots z of unity, $z^N = 1$).

We calculate $W(C)$ by expanding the integrand in expression (16.6) in powers of β_p. We choose for simplicity for the loop C a rectangle although the generalization to other contours is easy.

Any non-vanishing contribution must be invariant by the change of variables $U_\ell \mapsto z_\ell U_\ell$, where z_ℓ belongs to the centre. Let us consider one link belonging to the loop and multiply the corresponding link variable $\mathbf{U}(x, x+an_\mu)$ by z_0. We now multiply all link variables $\mathbf{U}(x+y, x+y+an_\mu)$, which are obtained by a translation y in the hyperplane perpendicular to n_μ, by z_y. Another link belonging to the loop belongs to the set but with opposite orientation. Plaquettes involving such variables involve them in pairs. For a result to be invariant and thus non-vanishing we need that the number of times each link variable appears in the direction n_μ minus the number of times it appears in the direction $-n_\mu$ vanishes (mod N). Thus,

we start adding plaquettes to satisfy this condition at point x. However, the addition of one plaquette does not change the total difference between the numbers of links in the $+n_\mu$ and $-n_\mu$ directions. Therefore, always at least one condition remains unsatisfied until the plaquettes reach the other link of the loop. We can then repeat the arguments for the remaining links of the loop and the new non integrated remaining links of the plaquettes. The number of required plaquette variables to get a non-vanishing result, is at least equal to the area of the rectangle, the minimal area surface having the loop as boundary. We can then perform the integrations which are just factorized group integrations. In this way, we get a contribution to $W(C)$ proportional to $(\beta_p)^A$, in which A is the number of plaquettes. The largest contribution corresponds to plaquettes covering the minimal area surfaces bounded by the loop. It is indeed obtained by covering the rectangle with plaquettes in such a way that each link variable appears only twice in either

orientation. For a rectangular loop $R \times T$, we just get

$$W(C) \sim \mathrm{e}^{RT \ln \beta_p} \,. \tag{16.12}$$

This results indicates that the potential between the static charges is linearly rising at large distance. Static charges creating the loop cannot simply be screened by the gauge field, in which case we would again get a perimeter law.

Remarks.

(i) If the centre is trivial, it is possible to form a tube along the loop and this implies a perimeter law. If, for example, the group is $SO(3)$, in the decomposition of a product of two spin 1 representations, we again find a spin 1 which can be coupled to a third spin 1 to form a scalar. Thus, two plaquettes can be glued to the same link of the loop without constraint on the orientation of the plaquette.

(ii) The asymptotic form (16.12) is also valid for the Abelian $U(1)$ lattice gauge theory. Therefore, in four dimensions, Wilson's loop has a perimeter law at any order in the weak coupling expansion and an area law at large coupling. We expect a phase transition between a low coupling Coulomb phase, described by a free field theory, and a strong coupling confined phase. This phase transition has been observed in numerical simulations. It seems to be first order, but this question has not been definitively settled. The existence of the transition is related to the compact nature of the $U(1)$ group which is only relevant on the lattice (lattice QED based on group elements is also called compact QED). Defects in which the group element varies by a multiple of 2π around a plaquette govern the dynamics of the transition. They correspond in the continuum to magnetic monopoles. In four dimensions monopole loops yield, for dimensional reasons, logarithmic contributions to the action, a situation reminiscent of the two-dimensional Coulomb gas. The separation of vortices in the Kosterlitz–Thouless (KT)

phase transition is here replaced by the separation of magnetic monopoles.

The string tension. The coefficient in front of the area is called the string tension σ,

$$\sigma(\beta_p) \underset{\beta_p \to 0}{\sim} -\ln \beta_p \, .$$

If no phase transition occurs when β_p varies from zero to infinity, the gauge theory leads to confinement. In this case, the behaviour of the string tension for β_p small is predicted by the renormalization group. Since σ has the dimension of a mass squared one finds

$$\sigma(e_0) \sim (e_0^2)^{-\beta_2/\beta_3^2} \exp\left(-1/\beta_2 e_0^2\right) . \tag{16.13}$$

in which e_0^2 is related to β_p by equation (16.9) and β_2, β_3 are two first coefficients of the RG β-function, the first being given in equation (15.25). A physical quantity relevant to the continuum limit can then be obtained

by dividing $\sqrt{\sigma}$ by its asymptotic behaviour. Let us define Λ_L as

$$\Lambda_L = a^{-1}(\beta_2 e_0^2)^{-\beta_3/2\beta_2^2} \exp\left(-1/2\beta_2 e_0^2\right),$$

then $\Lambda_L/\sqrt{\sigma}$ has a continuum limit. When one calculates σ by non-perturbative lattice methods, the verification of the scaling behaviour (16.13) indicates that the result is relevant to the continuum field theory and not only a lattice artifact.

It is possible to systematically expand σ in powers of β_p. The possibility of verifying that confinement is realized in the continuum limit, depends on the possibility of analytically continuing the strong coupling expansion up to the origin. Unfortunately, theoretical arguments lead to believe that, independently of the group, the string tension is affected by a singularity associated with the roughening transition, transition which, however, is not related to bulk properties. At strong coupling, the contributions to the string tension come from smooth surfaces. When e_0^2 decreases (β_p increases), one passes through a critical point e_{0R}^2, after which the relevant

surfaces become rough. At the singular coupling e^2_{0R}, the string tension does not vanish but has a weak singularity. Still at this point the strong coupling expansion diverges. Therefore, it is impossible to extrapolate to arbitrarily small coupling. The usefulness of the strong coupling expansion then depends on the position of the roughening transition with respect of the onset of weak coupling behaviour. Notice that numerically in the neighbourhood of the roughening transition, rotational symmetry is approximately restored (at least at large enough distance).

One can also calculate other quantities which are associated to bulk properties, and are, therefore, not affected by roughening singularities, such as the free energy (the connected vacuum amplitude) or the plaquette–plaquette correlation function. However, even for these quantities the extrapolation is not easy because the transition between strong and weak coupling behaviours is in general very sharp. This is confirmed by results coming from Monte Carlo simulations and is interpreted as indicating the

presence of singularities in the complex β_P plane close to the real axis. From the numerical point of view, it seems that the plaquette–plaquette correlation function is the most promising case for strong coupling expansion.

Remark. We note that the potential between static charges in the confined phase is linearly increasing in the same way as the Coulomb potential in one space dimension. This leads to the following physical picture: in QED the gauge field responsible of the potential has no charge and propagates essentially like a free field isotropically in all space directions. Conservation of flux on a sphere then yields the R^{2-d} force between the charges. However, in the non-Abelian case the attractive force between the gauge particles generates instead a flux tube between static charges in such a way that the force remains the same as in one space dimension.

Gauge symmetry breaking: Elitzur's theorem. Let us add a simple comment about the absence of a local order parameter in gauge theories. We have seen that in the temporal gauge the ground state is invariant under

space dependent gauge transformations. This property is incompatible with the existence of a local order parameter which is necessarily non-gauge invariant. Therefore, the question is: can a phase transition on the lattice lead to a spontaneous breaking of gauge invariance? To answer this qustion, we consider the transition probability at low temperature between two states, concentrated one around the minimal energy configuration $\mathbf{A}_\mu = 0$ and the other one around a pure non-trivial gauge. If the gauge function is different from zero only in a finite space volume, the cost in energy is the same as in a one-dimensional system and, therefore, the transition probability always remains finite independently of the number of space dimensions. Therefore, the quantum ground state is gauge invariant. Note that this argument does not apply to gauge transformations which do not vanish at large distances. Therefore, it does not forbid a spontaneous breaking of the global symmetry associated with the gauge group.

16.4 Mean field approximation

We have shown that the pure gauge lattice model yields at low temperature or coupling the continuum gauge theory. The continuum model allows, in perturbation theory, the separation of charges at large distances. On the contrary, at high temperature, charges are confined in the lattice model.

Therefore, it is necessary to investigate the possibility of phase transitions in lattice gauge theories. Since in the case of spin models, the mean field approximation gives a semi-quantitative understanding of the phase structure at least for $d > 2$, it is natural to also study gauge theories in the mean field approximation.

We introduce two sets of real matrices ϕ_ℓ and H_ℓ, in which the index ℓ stands for link. Then we write the partition function

$$\mathcal{Z} = \int \prod_\ell \mathrm{d}U_\ell \exp\left[-\beta_p \mathcal{S}(U)\right],$$

in which $\mathcal{S}(U)$ is the lattice action (16.5), as

$$\mathcal{Z} = \int \prod_{\text{links } \ell} \mathrm{d}\phi_\ell \, \mathrm{d}H_\ell \, \mathrm{d}U_\ell \exp \left[-\beta_p \mathcal{S}(\phi) + \sum_{links} \mathrm{tr}\, H_\ell \, (\phi_\ell - U_\ell) \right].$$

The introduction of the variables ϕ allows to express the action in terms of an average link variable. Since the average of an orthogonal (unitary) matrix is not orthogonal (resp. not unitary), we have defined ϕ as an arbitrary real (resp. complex) matrix. The variables H_ℓ represent directly at leading order the mean field which approximates the effect of the plaquette interaction.

The integral over the matrices U now factorizes into a product of integrals over each link variable:

$$\int \mathrm{d}U \ \mathrm{e}^{-\,\mathrm{tr}\, HU} = \mathrm{e}^{-\rho(H)},$$

in which $\rho(H)$ is thus a $G \times G$ group invariant function of H (H transforming under independent right and left multiplication).

The partition function becomes

$$\mathcal{Z} = \int \prod_\ell \mathrm{d}H_\ell \, \mathrm{d}\phi_\ell \exp\left\{ - \left[\beta_p \mathcal{S}(\phi) + \sum_\ell \big(\rho(H_\ell) - \operatorname{tr} H_\ell \phi_\ell \big) \right] \right\} .$$

We then look for saddle points in the variables H and ϕ. Since H and ϕ are general real or complex matrices, we expect to find many saddle points. However, both for simplicity and symmetry reasons, we look for solutions in which H_ℓ and ϕ_ℓ are constant on the lattice and multiple of the identity (up to a gauge transformation):

$$\phi_\ell = \varphi \mathrm{I} \,, \qquad H_\ell = h \mathrm{I} \,,$$

in which I is the identity matrix. Denoting by $\mathcal{S}(\varphi, h)$ the lattice action per link we then find

$$\mathcal{S}(h, \varphi) = \operatorname{tr} \mathrm{I} \left[-\tfrac{1}{2}(d-1)\beta_p \varphi^4 + V(h) - h\varphi \right] ,$$

in which we have defined $V(h)$ by

$$V(h) = \frac{\rho(h\mathrm{I})}{\operatorname{tr} \mathrm{I}}$$

with $V(h) = -\frac{1}{4}h^2 + O(h^4)$ for $SU(2)$. The saddle point equations are

$$\varphi = V'(h)\,, \quad h = -2(d-1)\beta_p \varphi^3.$$

We can eliminate φ and obtain

$$h = -2(d-1)\beta_p \left[V'(h)\right]^3. \tag{16.14}$$

For h small, $V'(h)$ is at least linear in h (as in $SU(2)$). We realize immediately the essential difference with the spin models we had considered so far. The right hand side of equation (16.14) is at least cubic in h instead of being linear. Thus, the equation has never a non-trivial solution arbitrarily close

to zero. For β_p small there exists only the trivial solution $h = 0$, which, according to the strong coupling or high temperature analysis, corresponds to the confined phase in which Wilson's loop follows an area law. For a critical value β_c, h jumps from zero to a finite value, indicating a first order phase transition. We recall that at a first order transition the correlation length, at least above the transition, remains finite. Therefore, the neighbourhood of the transition temperature does not define a continuum field theory, in contrast with the non-linear σ-model. Above β_c, the expectation value of Wilson's loop is given by

$$W(C) = \left\langle \mathrm{tr} \prod_{\text{all links}} \prod_{\ell \in C} \phi_\ell \right\rangle \sim \phi^{P(C)},$$

in which $P(C)$ is the perimeter of the loop. Therefore, Wilson's loop follows a perimeter law and the phase is deconfined. Above β_c we are in the low temperature phase which can be described by a continuum field theory and perturbation theory.

Discussion. Mean field theory is valid in high dimensions. Continuum field theory tells us that the zero temperature ($\beta_p = \infty$) is an IR stable fixed point for $d > 4$. Thus, the mean field result can only apply for $d > 4$. However, we would naively expect a second order phase transition in $4 + \varepsilon$ dimensions with a critical temperature of order ε, or $\beta_p \sim 1/\varepsilon$, in analogy with the non-linear σ-model. The open question is whether in any integer dimension $d > 4$ the transition is really second order.

For $d \leq 4$, the zero temperature is a UV fixed point. The simplest consistent scheme is one in which the critical temperature vanishes and the model always remains in the confined high temperature phase. The dimension $d = 4$ for gauge theories plays the role of the dimension $d = 2$ for the non-linear σ-model. The large momentum behaviour of correlation functions can be determined from perturbation theory, but no analytical method yields directly their low momentum behaviour and, therefore, for example, the spectrum of the theory. The only other analytical piece of

information available is the small coupling expansion in a finite volume of the eigenstates of the quantum Hamiltonian, which one can try to extrapolate by numerical methods towards the infinite volume limit using finite size scaling analysis. However, again there is numerical evidence of a sharp transition between the finite volume and infinite volume results, making the extrapolation difficult. The most promising quantities seem to be ratios of masses. This lack of reliable analytical methods explains the popularity of numerical simulations based on stochastic methods of Monte Carlo type in this problem.

Monte Carlo methods. We will not describe the increasingly sophisticated numerical methods which have been used in lattice gauge theories.

In pure gauge theories, the existence of phase transitions has been investigated for many lattice actions. For the gauge group $SU(3)$, relevant to the physics of Strong Interactions, the string tension has been carefully measured, the plaquette–plaquette correlation function has been studied to

determine the mass of low lying gluonium states. Finally, calculations have been performed at finite physical temperature, that is, on a $3+1$ dimensional lattice in the limit in which the size of the lattice remains finite in one dimension, this size being related to the temperature. In this way, the temperature of a deconfinement transition has been determined.

Fermions on the lattice. One important qualitative feature of Strong Interaction physics is the approximate spontaneous breaking of chiral symmetry (see section 12.5). However,non-trivial problems arise when one tries to construct a chiral invariant lattice action. One has the choice only between writing an action which is not explicitly chiral symmetric and in which one tries to restore chiral symmetry by adjusting the fermion mass term (Wilson's fermions), writing a chiral symmetric action with too many fermions (staggered or Kogut–Susskind fermions), or, as it has been more recently discovered various Dirac operators satisfying the Ginsparg–Wilson relation In the latter solution several implementations can be interpreted as adding

for the fermions an extra space dimension, which increases the already difficult computer problem. Indeed, an important practical difficulty also arises with fermions: because it is impossible to simulate numerically fermions, it is necessary to integrate over fermions explicitly. This generates an effective gauge field action which contains a contribution proportional to the fermion determinant and is, therefore, no longer local. The speed of numerical methods crucially depends on the locality of the action. This explains that most numerical simulations with fermions have been for some time performed in the so-called quenched approximation in which the determinant is neglected. This approximation corresponds to the neglect of all fermion loops and bears some similarity with the eikonal approximation. In this approximation, the approximate spontaneous breaking of chiral symmetry has been verified by measuring the decrease of the pion mass for decreasing quark masses. Owing to the difficulty of the problem, the numerical study of the effect of dynamical fermions at realistic lattice sizes and close enough

to the chiral limit has started in more recent years. The physical spectrum of hadrons can now be reproduced.