LATTICE GAUGE FIELDS

J.M. DROUFFE and C. ITZYKSON

Commissariat a L'Énergie Atomique, Division de la Physique, Service de Physique Théorique, CEN, Saclay, Boite Postale No2, 91190 Gif-s/-Yvette, France

Contents:

LATTICE GAUGE FIELDS

J.M. DROUFFE and C. ITZYKSON

*Commissariat a L'Énergie Atomique, Division de la Physique,
Service de Physique Théorique, CEN, Saclay,
Boite Postale No 2, 91190 Gif-s/-Yvette, France*

1. Introduction

In the last few years, new hopes have been raised in hadronic physics. For the first time, it seems that almost all the pieces of a coherent model for strong interactions are gathered and can be assembled. Hadrons behave as bound states of fundamental quark fields. These quarks belong to the fundamental representation of an SU(3) flavor symmetry and this allows the well known classification of hadrons following the eightfold way of Gell-Man and Néeman. The narrow resonances recently discovered can be included in this scheme by an extension to a larger symmetry group SU(4). The existence of a fourth quark had been predicted from hadronic weak interactions. There exists further experimental evidence for a composite hadronic structure. The deep inelastic scattering experiments indeed show scaling properties which lead to the parton picture of the hadronic matter. This seems to receive more support as one probes deeper and deeper with the available energies.

We have now a promising predictive model of weak and electromagnetic interactions based on a Yang–Mills Lagrangian [1] for the intermediate boson and the photon. For strong interactions also, a dynamic scheme can be derived using the same theoretical tool associated with a "gluon" field. The motivation in this case comes from the feature of small distance asymptotic freedom [2] of such an interaction. This leads to the colored quark model [3]. Here, the elementary fields are quarks and gluons. The ordinary quarks (carrying SU(3) or SU(4) flavor indices) are triplicated according to color and interact via an octet of colored gluons through the minimal Yang–Mills interaction. There is an exact symmetry with respect to the SU(3) color gauge transformations. Hadrons are color singlet bound states in this model. The short distance behaviour of Green functions is well understood in this context and can be studied using renormalization group techniques. Thanks to asymptotic freedom, the perturbation theory allows explicit and reliable calculation of short distance properties.

Large distance behaviour remains an unsolved problem due to infrared singularities. It is a controversial matter to see in perturbation theory whether or not quarks and gluons are prevented to appear as free, observable particles, whereas up to now only uncolored bound states are observed. One suspects the possibility that long range forces might permanently confine quarks and gluons within physical hadrons. The outstanding question is therefore to understand quark confinement and conciliate the apparently contradictory aspects of asymptotic freedom and infrared slavery.

The lattice gauge theories has been introduced by Wilson [4,5] as an attempt to answer these questions. Space, or space-time, is discretized by introducing an underlying lattice. This technique must be considered at this point as a theoretical aid or a computational method rather than a fundamental feature.

The aim of the present report is to review the achievements and the (numerous) unsolved problems of lattice gauge theories. Whether this will at least provide a reliable mean of computing the hadronic spectrum and its properties is by no means sure. Nevertheless they deserve interest for at least two reasons. The first has to do with the stimulus given to the search of a hadronic fully gauge invariant phase. The second is that they enlarge the set of lattice models with interesting new properties. They might even be used for different purposes as particles physics.

The models obtained in this way have some similarity with those of statistical mechanics. The existence of a phase transition is possible [4,6] in high enough dimension. A "quark phase", where free quarks can be observed, may exist for sufficiently low values of the couplings and can be treated in the framework of the conventional perturbation expansion of field theory. A "baryon phase" where only color singlets can propagate freely, arises in the high coupling region. This phase can be explored in lattice theories by high coupling expansions [7], similar to the high temperature series of statistical

mechanics. It is possible for instance to make the preliminary step towards a derivation of the meson or baryon mass spectrum [8].

A further step is to remove the artificial lattice in order to restore the continuous Lorentz invariant theory. One hopes that all continuous properties (and particularly the short distance properties) will reappear as the lattice spacing shrinks to zero in the vicinity of the transition point. Renormalization group techniques [9] provide a means to solve this problem and propose approximate recursion relations [10] for studying this limit.

Other techniques are presently under investigation in order to study the possibility of confinement and it is too early to decide which method is more efficient at this stage.

We have tried to give as self contained a treatment as possible. In section 2, we discuss the problems related with the construction of a lattice action which reduces to Yang–Mills minimal coupling Lagrangian in the continuous limit. An alternative formulation based on the Hamiltonian formalism is also derived.

Section 3 is devoted to the perturbative expansions for strong couplings where we summarize the different techniques and the corresponding results. In section 4, we gathered some non perturbative information on critical properties of these models. These two sections are more or less independent; nevertheless, some exact results obtained in section 4 are already used in the discussion of the perturbative results of section 3.

Problems related with the continuous limit are finally discussed in section 5. We concentrate mainly on the recursion relations provided by renormalization group techniques.

Deficiencies in our treatment of the subject can be partly remedied using the list of references which we tried to prepare as complete as possible up to beginning of 1977. We apologize for unvoluntary omissions.

2. Formulation

2.1. Generalities

As well known, continuous field theory is plagued by ultraviolet divergences cured perturbatively by renormalization. To study large distance properties and to escape small coupling expansions, a space time discretization acting as an ultraviolet cut-off is a useful device.

There exist several possibilities for introducing a lattice. We often (but not always) use in this paper a discretization over all (spatial and time) dimensions. The Minkowskian space is first transformed through a Wick rotation into an Euclidean space. All calculations are done in the framework of the Euclidean metric and the results are finally reinterpreted in the usual physical space just at the end of the procedure. No information is lost in this process; for instance, although the Wick rotation replaces the physical evolution according to the operator $\exp(-iHt)$ by an exponential decrease in $\exp(-Ht)$ of correlations, the physical energy spectrum can be recovered using appropriate techniques.

In this Euclidean space of dimension d, we introduce a hypercubical lattice with spacing a. This choice is not essential, and any other lattice, even not translationally invariant, may be considered. A particular choice does not destroy or modify the important features of the theory. The hypercubical lattice appears as the simplest one and applies to any dimension. It enjoys remarkably simple geometrical properties valuable in the explicit calculations, especially when using the renormalization

group techniques; furthermore it allows explicit resolution for some simple cases.

Dynamical quantities, such as the Lagrangian, are now involving fields defined on the lattice nodes. To construct it, we require that the usual quantities would be retrieved in the continuous limit, i.e. as the lattice spacing a goes to zero.

The quantization procedure (which cannot be performed following the usual canonical manner due to the Wick rotation) is achieved using the Feynman–Kac path integral formulation. One assumes the existence of an action integral $S[\phi]$ depending upon classical fields collectively denoted by $\phi(x)$. A physical quantity, as for instance the mean value $\langle X \rangle$ of a product of fields X is given by

$$\langle X \rangle = Z^{-1} \int \mathcal{D}\phi \, X \exp(S[\phi]) \tag{2.1.1}$$

$$Z = \int \mathcal{D}\phi \, \exp(S[\phi]). \tag{2.1.2}$$

In field theory, the formula is not a priori well defined. One has indeed to select precisely the functional space of classical fields and the integration procedure (unless one limits oneself to perturbation theory given in terms of gaussian integrals). However these problems do not arise in the lattice theory. The functional space is the (denumerable) Cartesian product of spaces which carry the values of field ϕ at each lattice site, and one integrates over these values. One may eventually introduce a space-time volume cut-off by using a finite lattice, with or without periodic boundary conditions. For boson fields, one has to choose a measure $\mathcal{D}\phi$ on the values of ϕ. For fermions, ϕ is considered as an element of a anticommuting Grassman algebra for which integration also makes sense [11]. Thus formula (2.1.1) has a precise mathematical sense and can be directly used in calculations.

The Lagrangian formalism on a Euclidean lattice keeps some flavour of the original Lorentz symmetry, that is the discrete symmetry interchanging time and spatial axis. The number of terms in any formula is then reduced. Furthermore the direct use of eq. (2.1.1) allows to carry systematically and easily at any order the diagrammatic high coupling expansion.

There exists in fact an alternative approach. Space-time remains Minkovskian and only the spatial dimensions are discretized. One recovers the usual problem of continuous evolution in time of fields defined on a three dimensional lattice. It is then possible to quantize the theory in a canonical way and thus to use a Hamiltonian formalism [12]. Although this method lacks some of the advantages described above, some physical quantities, and especially the mass spectrum, are more directly accessible to computations. We shall devote a special paragraph to this approach. Let us finally mention that Bardeen et al. [13], using an infinite momentum frame, discretize only two dimensions corresponding to the physical degrees of freedom of a gauge field and keep the longitudinal and time variable continuous.

The Euclidean Lagrangian lattice formalism is essentially identical with its statistical counterpart except for a different vocabulary. The field action can be identified with the energy of a configuration while the vacuum functional integral is the partition function. We shall see that the coupling constant plays a role alike to temperature. Finally the requirement of recovering a continuous theory in the limit of zero spacing is the same as finding a critical domain in statistical mechanics. The main physical difference is that for most dynamical systems (but not all) the relevant dimension of statistical mechanics is three, while Euclidean space-time has dimension four. This is a fortunate circumstance because dimension four seems the lowest critical dimension of systems with local symmetries. This means that the critical coupling is zero, thus suggesting that calculations performed on the lattice and properly renormalized are relevant to the continuous perturbation approach to the same problem.

2.2. Lattice actions

Once a lattice has been introduced, one is faced with the problem of defining an action. The only imperative requirement is to recover in the zero spacing limit the continuous field theory action, at least formally. We want also to perserve as much as possible the symmetry properties of the continuous models. We delay up to the next section the discussion of global and local gauge invariance and we focus here on the treatment of field derivatives.

A derivative $\partial_\mu \Phi(x)$ is replaced by a sum over finite differences between different lattice sites. The simplest choice is to use the nearest neighbour difference,

$$\partial_\mu \Phi(x) \to \frac{1}{a} (\Phi_{x+a\hat{\mu}} - \Phi_x) \tag{2.2.1}$$

where $\hat{\mu}$ is the unit vector in direction μ. Let us consider the case of free fields. As the continuous Lagrangian is split into kinetic and potential parts, the lattice action will be separated into bilocal and local parts. A boson free field will generate, with the convention of eq. (2.2.1) an Ising type (or similar) model with an action written as:

$$\int d^d x (\tfrac{1}{2} \partial_\mu \Phi^\dagger \cdot \partial^\mu \Phi + \tfrac{1}{2} m^2 \Phi^\dagger \Phi) \to -\frac{a^{d-2}}{2} \sum_{(xy)} \Phi_x^\dagger \Phi_y + \left(\frac{m^2 a^d}{2} + d a^{d-2} \right) \sum_x \Phi_x^\dagger \Phi_x. \tag{2.2.2}$$

The notation $\langle xy \rangle$ means that the summation runs over all *oriented* pairs of neighbouring sites and Φ is an n-component, complex or real, scalar field in which case the symbol $\Phi^\dagger$ is relevant. The range of Φ still remains to be specified.

The same recipe when used blindly in the fermionic case leads to the replacement:

$$\int d^d x \bar{\psi} (\partial\!\!\!/ + m) \psi \to \frac{a^{d-1}}{2} \sum_{\langle xy \rangle} \bar{\psi}_x \gamma_{xy} \psi_y + m a^d \sum_x \bar{\psi}_x \psi_x \tag{2.2.3}$$

where again the summation runs over all oriented pairs of sites; γ_{xy} is the matrix $\gamma \cdot (y-x)/a$ (i.e. $\pm \gamma_\mu$) and γ is the set of Euclidean Hermitian Dirac matrices satisfying the anticommutation relations

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \tag{2.2.4}$$

However, one must be careful with such a naïve replacement which may introduce spurious states. Let us discuss this difficulty by comparing bosonic and fermionic actions in momentum space. Using the Fourier transform

$$\phi_x = \int \frac{d^d k}{(2\pi)^d} \exp\{i k \cdot x\} \phi(k), \tag{2.2.5}$$

where the integral runs over one Brillouin zone of the reciprocal lattice ($|k_\mu| \leq \pi/a$ for instance), the action (2.2.2) takes the following form

$$\tfrac{1}{2} \int \frac{d^d k}{(2\pi)^d} \left(\sum_\mu \frac{4}{a^2} \sin^2 \left(\frac{a k_\mu}{2} \right) + m^2 \right) \tilde{\Phi}^\dagger(-k) \tilde{\Phi}(k). \tag{2.2.6}$$

As a result, each mode contributes to the action a quantity $(m^2 + (4/a^2)\sum_\mu \sin^2(\tfrac{1}{2} a k_\mu))$ rather than the standard form $(m^2 + k^2)$ (see fig. 1). Nevertheless, the two curves have only one minimum in the Brillouin zone at $k = 0$, at which point they coincide up to $O(a^2 k^4)$ terms. This ensures the correct continuous limit. Let us mention that it is possible to obtain the classical shape in $k^2 + m^2$ (in the first
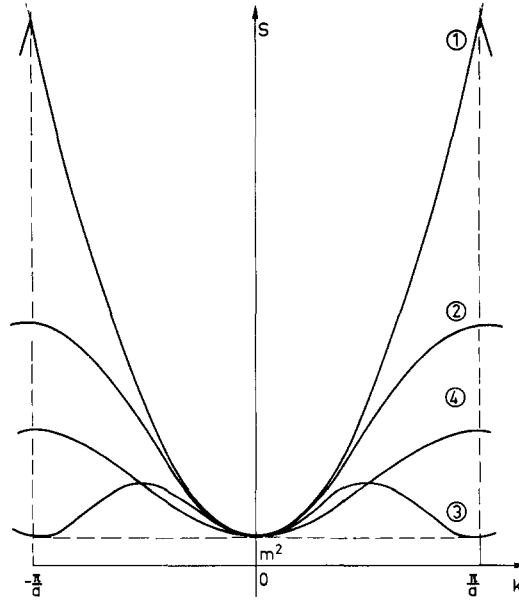
Fig. 1. Sketch of different lattice actions in the first Brillouin zone. 1) classical shape $k^2 + m^2$ of eq. (2.2.7), 2) bosons (2.2.2), 3) fermions with eq. (2.2.3), 4) Wilson prescription for fermions, eq. (2.2.9).

Brillouin zone) if one renounces to the nearest neighbour prescription [14]. One has indeed to introduce long range interactions

$$\partial_\mu \Phi(x) \rightarrow \int \frac{d^d k}{(2\pi)^d} ik_\mu \tilde{\Phi}(k) \exp\{ik \cdot x\} = \sum_y \left[\int \frac{d^d k}{(2\pi)^d} ik_\mu \exp\{ik \cdot (x - y)\} \right] \Phi_y. \tag{2.2.7}$$

Such ideas might be useful [15] when one is faced with the problem of restoring Lorentz or chiral invariance which has been destroyed by the introduction of the lattice.

Let us now turn to the fermionic case. Using the Fourier transform (2.2.5), we can rewrite the free field action (2.2.3) as

$$\bar{\psi}(-k) \left[i \sum_\mu \gamma_\mu \frac{\sin ak_\mu}{a} + m \right] \psi(k) \tag{2.2.8}$$

and thus the action behaves as $m^2 + (1/a^2)\Sigma_\mu \sin^2 ak_\mu$. There are now two equal minima in a Brillouin zone. One is located around $k = 0$ and gives the correct continuous limit. However other modes around $k = \pm\pi/a$ survive, they will carry infinite momentum in the continuous limit and might be nevertheless excited. This unpleasant phenomenon must be suppressed if one wants to obtain a reasonable model for field theory. The above recipe of eq. (2.2.7) may be used. Nevertheless, the nearest neighbour prescription can be retained with some modifications in the action (2.2.3). Wilson [16] proposes to use $1 + \gamma_{xy}$ rather than γ_{xy} and to write

$$\int d^d x \bar{\psi}(\partial + m)\psi \rightarrow \frac{a^{d-1}}{2} \sum_{\langle xy \rangle} \bar{\psi}_x (1 + \gamma_{xy})\psi_y + (ma^d - da^{d-1}) \sum_x \bar{\psi}_x \psi_x. \tag{2.2.9}$$

This modification adds a supplementary term best appreciated in momentum space, where the
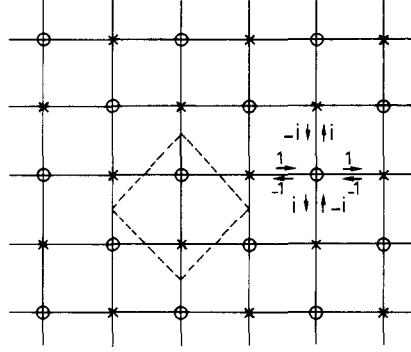
Fig. 2. Partition of sites; encircled nodes carry upper spinor components, crossed nodes lower components. The new elementary cell and the numbers γ_{xy} are shown.

action now reads:

$$\bar{\psi}(-k)\left[i \sum \gamma \frac{\sin ak}{a} + m + \sum \frac{\cos ka - 1}{a}\right]\psi(k). \tag{2.2.10}$$

This increases the secondary minima without influencing the small ka behaviour (see fig. 1) which will be the only one to (hopefully) survive in the continuous limit. This prescription is however arbitrary to a large extent and does not fully exploit the above degeneracy.

Alternatively, Casher and Susskind [17] suggest to compensate the duplication of modes by introducing fewer components of the Fermi fields on each site and increasing correspondingly the size of the basic cell of the lattice. Let us illustrate this method with a two-dimensional example. Attached to each site there is a one component anticommuting field. The usual two-component spinor is recovered in the continuous limit by gathering two neighbouring fields (see fig. 2). Therefore the lattice is partitioned into two sets of sites (call them even and odd), each of which carries upper or lower spinor components. The original action (2.2.3) had four minima within a Brillouin zone $|k_\mu| \le \pi/a$. Since in configuration space the basic cell has increased by a factor $\sqrt{2}$ in size, the corresponding one on the reciprocal lattice contains now only two minima, a count which now agrees with the number of degrees of freedom. A realization of this idea is achieved in this two-dimensional case with the action

$$\frac{a^{d-1}}{2}\sum_{\langle xy \rangle} \chi_x \gamma_{xy} \chi_y + ma^d \sum_x \bar{\chi}_x \chi_x. \tag{2.2.11}$$

The two anticommuting fields χ_{odd} and χ_{even} of a cell give in the continuous limit the spinor $\psi = \binom{\chi_{\mathrm{odd}}}{\chi_{\mathrm{even}}}$. With the choice of the numbers γ_{ij} indicated in fig. 2, one recovers the usual continuous action (with the realization of the Clifford algebra $\gamma_1 = \sigma_1 \equiv \binom{0\ 1}{1\ 0}$, $\gamma_2 = \sigma_2 \equiv \binom{0\ i}{-i\ 0}$.

2.3. Global and local invariance

We now assume the theory to be invariant under the transformation of a (compact) symmetry group G. The fields (denoted collectively by Φ_x^α) belong to a representation D of this group, i.e. they transform as

$$\Phi^\alpha \to D(\mathcal{R})^\alpha_\beta \Phi_x^\beta, \qquad \mathcal{R} \in G. \tag{2.3.1}$$

The action is invariant under G. For instance, the typical bilocal interaction is written as

$$\sum_{\langle xy \rangle} \bar{\Phi}_x \Phi_y = \sum_{\langle xy \rangle} \bar{\Phi}_x^\alpha g_{\alpha\beta} \Phi_y^\beta \qquad (2.3.2)$$

$\bar{\Phi}_i^\alpha$ is a short hand notation for $\varphi_i^{\alpha\dagger}$ (bosons) or $\bar{\psi}_i^\alpha$ (fermions), which transforms under the adjoint representation of D, and $g_{\alpha\beta}$ are the coefficients of a quadratic form invariant under D.

If we allow the transformation R to depend on the lattice site, the bilocal part of the action does not remain invariant. As in the continuous case [18], our goal is now to introduce extra degrees of freedom in order to implement such a local invariance. A new field R_{xy}, attached to each oriented link $\langle xy \rangle$ and taking its value in the group G, is introduced.

The finite difference $\Phi_y^\alpha - \Phi_x^\alpha$, which replaces in the discretization procedure the derivative $\partial_\mu \Phi(x)$, is replaced by the *covariant difference*:

$$D(R_{xy})_\beta^\alpha \Phi_y^\beta - \Phi_x^\alpha. \qquad (2.3.3)$$

We require that this new quantity has the same transformation law as Φ_x^α (i.e. eq. (2.3.1)) in order to ensure the local invariance of the action. This is achieved with the following behaviour of the fields

$$\Phi^\alpha \to D(\mathcal{R}_x)_\beta^\alpha \Phi_x^\beta, \qquad R_{xy} \to \mathcal{R}_x R_{xy} \mathcal{R}_y^{-1}. \qquad (2.3.4)$$

Finally, hermiticity conditions connect gauge fields located on links differing only by their orientation

$$R_{yx} \equiv R_{xy}^{-1}. \qquad (2.3.5)$$

With these prescriptions, the bilocal interaction (2.3.2) reads

$$\sum_{\langle xy \rangle} \bar{\Phi}_x^\alpha D(R_{xy})_{\alpha\beta} \Phi_y^\beta. \qquad (2.3.6)$$

It may be useful at this stage to verify the compatibility between these prescriptions and the usual Yang–Mills model in the case of a continuous Lie group G. We have only in mind here to show that the naïve $a \to 0$ limit of the discretized action reproduces the minimal coupling prescription. We naturally assume that $R_{x,x+a\hat{\mu}}$ is close to the identity and write therefore

$$R_{x,x+a\hat{\mu}} = \exp(iga\mathcal{A}_\mu(x)), \qquad (2.3.7)$$

where $\mathcal{A}_\mu(x)$ belongs to the Lie algebra $\mathcal{g}$ of G. In the limit $a \to 0$, all other quantities being kept fixed:

$$\frac{1}{a}(D(R_{xy})\Phi_y - \Phi_x) \to [\partial_\mu + igD(\mathcal{A}_\mu(x))]\Phi(x). \qquad (2.3.8)$$

The right-hand side is identified with the usual covariant derivative.

Furthermore we also need in the action a term involving only gauge degrees of freedom. In order to preserve local invariance, one considers a closed curve $x_1 x_2 \ldots x_k x_1$ drawn on the lattice and computes

$$\beta\chi(R_{x_1 x_2} R_{x_2 x_3} \ldots R_{x_k x_1}) \qquad (2.3.9)$$

where χ is an arbitrary class function (character) on the group (thus (2.3.9) is independent of the choice of an origin x_1 on the curve). One has to sum over all closed curves equivalent with respect to the displacement group of the lattice. Although any closed curve is a good candidate, we choose the simplest one built on the sides of an elementary square of the lattice (the so-called "plaquette").

In general, these closed curves must be oriented. The orientation is however irrelevant if the character χ is real – i.e. χ is the character of a representation equivalent to its adjoint –. The same remark is valid for links.

In the continuous limit, the gauge term generates the usual kinetic part of the gauge field $-\frac{1}{4}\chi(F_{\mu\nu}^2)$, with $F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + g[\mathcal{A}_\mu, \mathcal{A}_\nu]$. Indeed, as can be shown after some straightforward algebra

$$R_{x,x+a\hat{\mu}}R_{x+a\hat{\mu},x+a\hat{\mu}+a\hat{\nu}}R_{x+a\hat{\mu}+a\hat{\nu},x+a\hat{\nu}}R_{x+a\hat{\nu},x} = \exp[ga^2(F_{\mu\nu} + O(a))]. \tag{2.3.10}$$

Expanding the exponential and taking the character gives first a constant term, then a linear term, cancelled by the contribution of the oppositely oriented plaquette. Finally, the first, non trivial, non vanishing term is just $-\frac{1}{4}\chi(F_{\mu\nu}^2)$, provided that

$$\beta = a^{d-4}/2g^2. \tag{2.3.11}$$

All higher order terms vanish in the $a \to 0$ limit. Equation (2.3.11) shows that the strong coupling region is governed by small values of β (i.e. large temperatures).

The final form of the invariant action reads in a short hand notation

$$S[R, \Phi] = \alpha \sum_{\langle xy \rangle} \bar{\Phi}_x R_{xy} \Phi_y + \beta \sum_P \chi(RRRR) + \sum_x V(\Phi_x) \tag{2.3.12}$$

and, in particular, Wilson's prescription for fermions leads to

$$S[R, \psi] = \frac{a^{d-1}}{2} \sum_{\langle xy \rangle} \bar{\psi}_x(1 + \gamma_{xy})R_{xy}\psi_y + \frac{a^{d-4}}{2g^2} \sum \chi(RRRR) + \sum_x (ma^d - da^{d-1})\bar{\psi}_x\psi_x. \tag{2.3.13}$$

Using this action with appropriate boundary conditions, only gauge invariant terms have non vanishing expectation values. In other words, the lattice formulation has bypassed the problem of the invariant quantization occuring in the continuous case (with its arbitrary choice of auxiliary conditions, Fadeev–Popov ghosts,...). This is related to the fact that, within a finite volume of space-time, the gauge freedom only contributes a finite factor proportional to a power of the volume of the (compact) group. This invariance property however does not ensure confinement, which results of the vanishing contribution of large matter field loops as we shall see later on.

The group structure – in particular whether it is a finite, discrete or continuous group – does not play any particular role in the above construction. Nor does the distinction between Abelian or non Abelian case. These will come into play in the discussion of critical properties of these models.

The usual characterization of systems according to the nature of their order parameter is lost at first sight when dealing with locally invariant systems. In the usual description, the order parameter appears as a local average (as for instance the magnetization or the gas density in a lattice gas model). Here, since only gauge invariant concepts are available to the analysis, a more refined instrument, the Wilson loop average, will be required to distinguish between phases of the system. This quantity generalises the usual correlation function

$$\langle \phi^{\dagger\alpha}(x)\phi^\beta(y) \rangle. \tag{2.3.14}$$

As was done for eq. (2.3.9), one assigns to each closed path C on the lattice the mean value

$$W(C) = \left\langle \chi\left(\prod_{\langle xy \rangle \in C} R_{xy} \right) \right\rangle. \tag{2.3.15}$$

The behaviour of (2.3.14) for large loops C enables one to detect the transition between ordered and

disordered phases and characterizes a locally invariant system. This discussion will be carried in detail in section 4.

2.4. Hamiltonian formalism

The lattice gauge theory can also be studied in the Hamiltonian formulation initiated by Kogut and Susskind [12] using only a spatial lattice. This can be recovered starting from the Lagrangian formalism and taking the limit of zero lattice spacing for a timelike direction [19] – by time, we mean that a preferred direction has been singled out on the lattice – . In this limit, the Hamiltonian will be identified with the logarithm of the transfer matrix which we now define.

The transfer matrix connects field configurations at successive times. Its matrix elements $T(\phi_{t+1}, \phi_t)$ are obtained by factoring out in $\exp(S[\phi])$ the contribution involving the interaction between fields at time t and $t + 1$, and, say, half of the piece involving only fields at time t or $t + 1$. The last part of the prescription is to some extent arbitrary. All that is required is that the partition function can be written in terms of T,

$$Z = \operatorname{tr} T^L = \int \ldots \mathcal{D}\phi_{t+1} \mathcal{D}\phi_t \ldots T(\phi_{t+1}, \phi_t) T(\phi_t, \phi_{t-1}) \ldots, \tag{2.4.1}$$

using appropriate (cyclic) boundary conditions. The transfer matrix plays the role of a unit (Euclidean) time evolution operator. Its logarithm can therefore be identified with the Hamiltonian.

Let us apply this formalism to the gauge field action. Local gauge invariance allows one to fix R_{xy} for an arbitrary set of links $\langle xy \rangle$, provided this set does not contain any closed loop. In particular, this applies to all timelike links where R_{xy} can be chosen equal to unity. This choice will be called Coulomb gauge and is necessary to our purpose. Let us use a special notation a_0 for the timelike lattice spacing. The action for a pure gauge field in the Coulomb gauge then reads

$$S[R] = \frac{a_0 a^{d-5}}{2g^2} \sum_{\substack{\text{spatial} \\ p}} \chi(RRRR) + \frac{a_0^{-1} a^{d-3}}{2g^2} \sum_{\substack{\text{spatial} \\ \langle xy \rangle}} \chi(R_{xy} R^{-1}_{x+a_0\hat{t}, y+a_0\hat{t}}). \tag{2.4.2}$$

The transfer matrix operates on a Hilbert space of configurations which is a direct product of spaces of square integrable functions over the group G. An (orthogonal but non normalizable) basis $|\{R\}\rangle$ is characterized by the set of values R_{xy} for each link of a $(d-1)$-dimensional space lattice. In this basis, the transfer matrix elements read

$$\langle\{R'\}|T|\{R\}\rangle = \exp\left\{\frac{a_0 a^{d-5}}{4g^2} \sum_p \left[\chi(RRRR) + \chi(R'R'R'R')\right] + \frac{a^{d-3}}{2g^2 a_0} \sum_{\langle xy \rangle} \chi(R_{xy} R'^{-1}_{xy})\right\}. \tag{2.4.3}$$

If $\mathcal{R}$ denotes a set of group elements $\mathcal{R}_{xy}$ for each link, one considers the direct product over the links of the regular representation of the gauge group, which acts on the space as

$$\hat{U}(\mathcal{R})|\{R\}\rangle = |\{\mathcal{R}R\}\rangle. \tag{2.4.4}$$

The introduction of the set of operators $\hat{R}^{r;\alpha\beta}_{xy}$ defined by

$$\hat{R}^{r;\alpha\beta}_{xy}|\{R\}\rangle = D^r(R_{xy})^{\alpha\beta}|\{R\}\rangle \tag{2.4.5}$$

allows to write the transfer matrix in operatori form as

$$T = \int \prod_{\langle xy \rangle} \mathcal{D}\mathcal{R}_{xy} \exp\left\{\frac{a^{d-3}}{2g^2 a_0} \sum_{\langle xy \rangle} \chi(\mathcal{R}_{xy})\right\} \hat{U}(\mathcal{R}) \exp\left\{\frac{a_0 a^{d-5}}{4g^2} \sum_p \chi(\hat{R}\hat{R}\hat{R}\hat{R})\right\}. \tag{2.4.6}$$

The limit $a_0 \to 0$ is obtained after some albebra; in this limit, only a neighbourhood of unity contributes in the integral over $\mathfrak{R}$. A parametrization of $\mathfrak{R}$ in the Lie algebra leads to gaussian integrations which can then be performed. The final result is the Hamiltonian (up to an (infinite but unsignificant) constant term):

$$H = \lim_{a_0 \to 0} \frac{1}{a_0} \text{Log } T = \frac{g^2}{a^{d-3}} \sum_{\langle xy \rangle} \sum_i \frac{\hat{l}_{xy}^{i^2}}{\chi(l_{xy}^{i^2})} + \frac{a^{d-5}}{2g^2} \sum_p \chi(\hat{R}\hat{R}\hat{R}\hat{R}) \tag{2.4.7}$$

when we have used an orthogonal basis l_i of the Lie algebra with respect to the Cartan–Killing metric. In the first term appears the quadratic Casimir operator (which is, a scalar in each irreducible subspace).

The two types of terms in (2.4.7) represent the contribution of "vertical" and "horizontal" (i.e. fixed time) plaquettes. The first one may be identified as an electric, and the second as a magnetic term, very much as in the Coulomb formulation of quantized electrodynamics. Alternatively, with the vector potential and the electric field playing the role of canonically conjugated variables, the first terms represent kinetic contributions, and the second potential energy.

The space on which the transfer matrix acts can be considered as a Fock space. We now use for each link a basis $|rm\rangle$ for the irreducible subspace associated to the representation r. Such a state is called a "string bit" if r is not the trivial representation. The vacuum $|0\rangle$ is the state without any string bits, i.e. with all links associated to the trivial representation. In this picture, the operators $\hat{R}_{xy}^r$ map the vacuum on a state with one string bit, located on the link $\langle xy \rangle$ and carrying the representation index r. This basis diagonalizes the first term of the Hamiltonian (2.1.7), which is proportional to the quadratic Casimir operator. Thus it is well adapted to a description of states in the strong coupling region, where the second term in the Hamiltonian becomes negligible. A topological description of gauge invariant states will be given in the next subsection in the case of Wilson's model.

The same procedure may be used with the complete model including matter fields. For instance, the action (2.3.13) for fermions with Wilson's prescription leads to the Hamiltonian [19]:

$$H = \frac{g^2}{a^{d-3}} \sum_{\langle xy \rangle} \sum_i \frac{\hat{l}_{xy}^{i^2}}{\chi(l_{xy}^{i^2})} + \sum_x (ma^{d-1} - (d-1)a^{d-2})\hat{\bar{\psi}}_x \hat{\psi}_x$$

$$+ \frac{a^{d-2}}{2} \sum_{\langle xy \rangle} \hat{\bar{\psi}}_x (1 + \gamma_{xy}) \hat{R}_{xy} \hat{\psi}_y + \frac{a^{d-5}}{2g^2} \sum_p \chi(\hat{R}\hat{R}\hat{R}\hat{R}). \tag{2.4.8}$$

The choice between the two formulations depends on the goal of the investigation. The Lagrangian formalism seems to be well adapted to the study of transition points and the calculation of Green functions while the determination of the energy spectrum is carried over by using the Rayleigh Schrödinger perturbation theory in the Hamiltonian formalism.

2.5. Some specific models

In this subsection, we present some models which allows to test strong coupling expansions in the lattice theory. Finally we describe Wilson's tentative model for strong interactions.

In statistical mechanics, the Heisenberg model [20] has been extensively studied. It is in fact the lattice version of the non-linear σ-model where a n-component bosonic field k is submitted to the supplementary condition $k^2 = 1$. The corresponding action

$$S[k] = \alpha \sum_{\langle xy \rangle} k_x \cdot k_y \tag{2.5.1}$$

is invariant under a global $O(n)$ symmetry. In particular, the case $n = 1$, which corresponds to the discrete symmetry group Z_2 (Ising model), is exactly soluble in dimensions 1 and 2. Similar results are valid for the corresponding gauge model [21]

$$S[k, R] = \alpha \sum_{\langle xy \rangle} k_x R_{xy} k_y + \beta \sum_p \chi(RRRR). \tag{2.5.2}$$

In particle physics, similar models with continuous symmetries have been studied in the context of the Higgs mechanism of spontaneous symmetry breaking [21]. The gauge field acquires a non zero mass while the $(n - 1)$ massless Goldstone bosons are expected to disappear. These features can be retrieved in the lattice case. We choose a particular direction K and perform the local gauge transformation given by $k_x = \mathcal{R}_x k$. With this choice of gauge, the action only involves the gauge degrees of freedom

$$S[R] = \beta \sum_p \chi(RRRR) + \alpha \sum_{\langle xy \rangle} K R_{xy} K \tag{2.5.3}$$

and the bosonic fields have disappeared. The parameter β is related to the coupling constant g, and α to the squared mass $m^2 = \alpha g^2 a^{2-d}$ of the gauge field.

Two models involving fermions are soluble in the continuous limit. It is thus interesting to compare them to their discrete counterparts. Schwinger's two-dimensional massless electrodynamics [23–27] exhibits no free fermions or photons, but neutral bound states of mass $e/\sqrt{\pi}$. This model is indeed a good laboratory for testing lattice models which naturally exhibit this confinement property.

The resummation method of t'Hooft [28–30] allows to show the same phenomenon for $SU(N)$ two-dimensional gauge models in the large N limit. Planar diagrams similar to those of the dual string model are dominant in this limit and it would be interesting to compare them [27,31] to the strong coupling expansion diagrams of the lattice theory.

Finally, we describe a more realistic model proposed by Wilson [16] as a tentative scheme for strong interaction dynamics. The fundamental fields are quarks. Quarks carry the indices of the lowest dimensional representation of a flavor group. With an $SU(3)$ group, we have therefore three quarks (up, down and strange), but it is of course possible to use larger flavor groups (such as $SU(4)$ for instance, with a fourth charmed quark). The choice of the flavor group, although important for physical applications, is in fact unessential for our purpose and only multiplies the number of quarks. It need not be an exact symmetry; indeed it is not, and a possible breaking is achieved by giving different bare masses to the quarks.

One postulates a supplementary $SU(3)$ exact color symmetry [2]. Each quark is now triplicated in three colors and we require the symmetry to be satisfied locally. We are then led to introduce gauge fields and use the lattice action (2.3.13) or the Hamiltonian (2.4.8). As a consequence of the exact color symmetry, only color singlets can propagate and we are led to the following topological classification (see fig. 3 and appendix A).

(i) States with no fermions are 1 dimensional complexes built with string bits, for instance closed curves. The simplest one is the "boxciton", made of four string bits bounding an elementary square of the lattice. The eigenvalues of the $SU(3)$ Casimir operator are well known; a string bit carrying the fundamental representation 3 (or $\bar{3}$) contributes a factor $\frac{4}{3}$ to the first term of the Hamiltonian (strong coupling limit). Similarly, the representations 6 (or $\bar{6}$) and 8 lead to factors $\frac{10}{3}$ and 3, and so on. We give
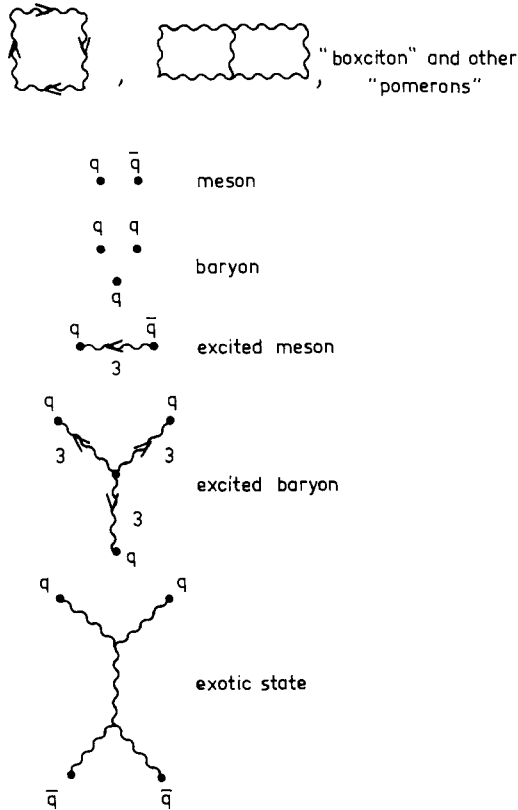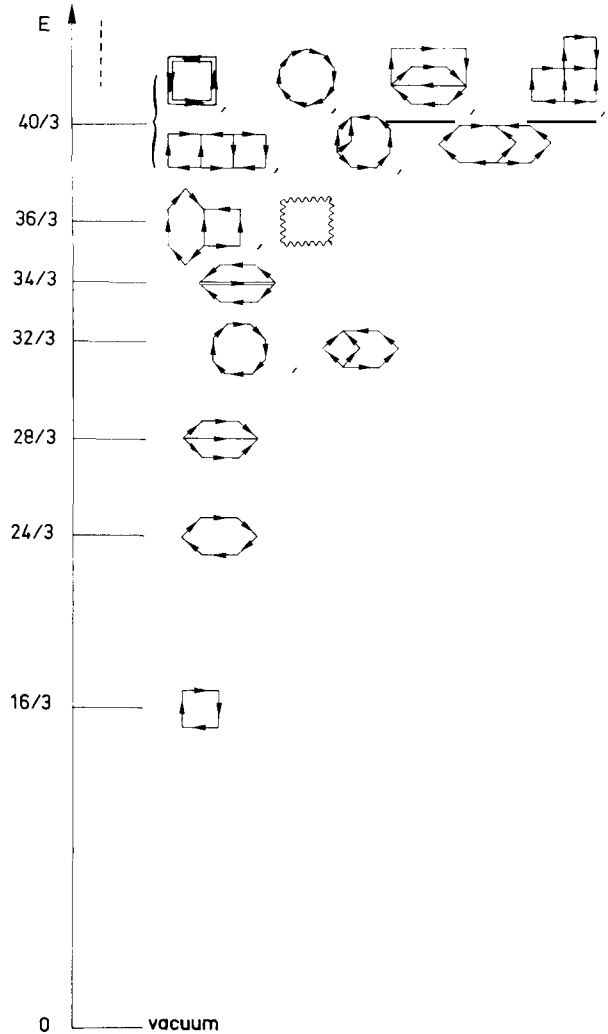
Fig. 3. Some possible states in Wilson's model.

Fig. 4. Spectrum of unperturbed pure SU(3) gauge field; →—: 3-string bit, ⇒==: 6-string bit, ⌇⌇: 8-string bit.

in fig. 4 the lowest levels of these states without quarks. They have the vacuum quantum numbers and could be identified with the phenomenological Pomeron (?).

(ii) Mesons are states with a quark–antiquark pair of the same color at the same site. The corresponding unperturbed masses are the sum of quark constituent masses. If one considers the spin $\frac{1}{2}$ of quarks together with the SU(3) flavor indices, these states are classified following an SU(6) scheme. We get here a singlet (η meson) and a 35-uplet (the octet of scalar mesons and the nonet of vector mesons).

(iii) Similarly, three quarks at the same site, antisymmetrized with respect to color indices, form the fundamental baryonic 56-uplet. (The usual spin $\frac{1}{2}$ octet and spin $\frac{3}{2}$ decuplet.)

(iv) Other states of higher energy involve quarks and string bits. Excited mesons or baryons contain a pair quark–antiquark, or three quarks, at different sites, connected by string bits. There exists also exotic states with more quarks.

This classification is of course valid only for infinite coupling constant g. A perturbative analysis of the corrections is constructed in the next section.

3. Strong coupling expansions

3.1. Fundamental features

In this section, we derive a systematic strong coupling expansion in the Euclidean Lagrangian formulation. We first focus on the integration over gauge degrees of freedom, in eq. (2.1.1). This integration can be performed for any gauge group and exhibits simple properties. The main feature is an expansion in terms of bidimensional diagrams with simple topological properties.

Consider the general action given in eq. (2.3.12). We are interested by the high coupling region in which α and β are small. Let us restrict ourselves at first to the case of a pure gauge field.

The exponentiated action is written as a product, each term of which corresponds to a definite plaquette

$$\exp[\beta\sum_p \chi(RRRR)] = \prod_p [\beta_0 + \sum_{r\neq 0} \beta_r \chi^r(RRRR)]. \tag{3.1.1}$$

Each factor on the left-hand side being a class function has been expanded on the right-hand side as a linear combination over the characters of all the irreducible representation r of the group, including the trivial one (referred as $r = 0$). The product runs over all *unoriented* plaquettes. In fact, each plaquette has been conventionally oriented: this allows to compute the product $RRRR$. Reversing the orientation replaces the representation r by its adjoint denoted $\bar{r}$. The contribution remains unchanged since $\beta_r = \beta_{\bar{r}}$. This is due to hermiticity or in other words, to the fact that the original χ in the action is chosen real.

As β approaches zero, all the coefficients β_r are at least of order β and vanish, except β_0 which tends to unity and thus plays a special role. Considering a lattice with N nodes and cyclic boundary conditions, we factorize our β_0 for each of the $\frac{1}{2}N d(d-1)$ plaquettes. The expansion of the product (3.1.1) can be interpreted in terms of *graphs*. A graph is a finite set of different plaquettes, to each of which a non trivial representation has been associated. A particular plaquette is assigned a factor $(\beta_r/\beta_0)\chi^r(RRRR)$. The contribution of a graph to $Z\beta_0^{-Nd(d-1)/2}$ is the product of all its constituent plaquette contributions integrated over the R fields. One has then to sum over all possible graphs.

Selection rules for the graphs are derived from the orthogonality properties of characters. This was indeed the reason for introducing such an expansion in the first place. One has the following identity:

$$\int \chi^r(RS)\chi^{r'}(R^{-1}T)\mathcal{D}R = \delta_{rr'}\chi^r(ST), \tag{3.1.2}$$

and, in particular

$$\int \chi^r(R)\mathcal{D}R = \delta_{r0}. \tag{3.1.3}$$

As a consequence, a plaquette cannot have free edges, and, if a link is shared by only two plaquettes, they must carry the same representation. Geometrically a diagram is thus a two-dimensional polyhedral complex without boundary lines (some elementary definitions of algebraic topology are recalled in appendix A). Each regular component carries the index of an irreducible representation r and

contributes a factor $(\beta_r/\beta_0)^A$ (A being its area, i.e. the number of component plaquettes) times the character χ^r of the product of gauge fields along its boundary. Integration over each interior link has indeed been performed but integration along the branch lines has still to be performed. This yields a factor proportional to the number of times the trivial representation occurs in the product of representations carried by the surfaces intersecting along this line. Let us mention that a particular regular component may be non-orientable provided it carries the index of a self-adjoint representation. The simplest examples are the closed surfaces. For instance, Z is given at lowest order by the cubic graph

$$Z = \beta_0^{Nd(d-1)/2}\left[1 + \frac{Nd(d-1)(d-2)}{6}\sum_{r\neq 0}(\beta_r/\beta_0)^6 + \ldots\right].$$ (3.1.4)

We investigate now the complete model with matter fields Φ. In the functional space generated by Φ, let us introduce a basis $U^{rm}(\Phi)$ for the irreducible subspace associated to the representation r. All representations do not occur; for instance, only totally symmetric representations are present in the case of scalar bosonic fields. In the fermionic case, they are in finite number (since the Grassman algebra has a finite dimension) and are totally antisymmetric for one-component spinors. It is now possible to write a formula analogous to (3.1.1) for the matter field interaction

$$\exp\left[\alpha\sum_{\langle xy\rangle}\bar{\Phi}_x R_{xy}\Phi_y\right] = \prod_{\langle xy\rangle}\left[\alpha_0 + \sum_{r\neq 0}\alpha_r P^r(\bar{\psi}_x R_{xy}\Phi_y)\right].$$ (3.1.5)

The zonal spherical functions P^r associated to the representation r are defined as

$$P^r(\bar{\Phi}R\Phi) = U_m^r(\bar{\Phi})D^r(R)_{m'}^m, U^{rm'}(\Phi).$$ (3.1.6)

Graphs are now composed of a 2-dimensional complex (pure gauge field interactions) and a 1-dimensional complex (matter field interactions). This 1-dimensional complex must be part of the contour of the plaquette set, due to the identity

$$\int P^r(\bar{\Phi}R\Phi)\mathcal{D}R = 0 \quad \text{if} \quad r \neq 0.$$ (3.1.7)

Boundary lines are now permitted; surfaces must be bounded by links carrying the same representation index since

$$\int \chi^r(R^{-1}S)P^{r'}(\bar{\Phi}R\Phi)\mathcal{D}R = \delta_{rr'}P^r(\bar{\Phi}S\Phi).$$ (3.1.8)

3.2. Physical consequences

Before embarking into the specifics of the numerical computations, let us outline the physical outcome. The main result is quark confinement in the strong coupling region. We already know that only singlets for the local gauge group may have non zero mean values. In particular, any Green function for separate quarks vanishes. Nevertheless, the preceding argument must be refined since it may remain possible to observe a quark (its energy, its flavor state, ..., except its color) with some local detector. We now carry a more careful analysis and show that the energy for separating a pair quark–antiquark (without creating new pairs) grows linearly with the separation distance. This is indeed the signal of confinement. Let us consider the following process: a pair is produced at some

point and is separated at a certain distance r. The quarks propagate in this situation during some time t, then annihilate (see fig. 5). In the strong coupling scheme, the lowest order diagram consists of the minimal surface made of plaquettes and bounded with the quark loop. There are $r\,t$ plaquettes at least (distance and time are given in lattice spacing units). Hence the contribution is $\exp\,[r\,t\,\text{Ln}\beta + 2(r+t)\text{Ln}\,\alpha]$. As the probability of such a process is expressed in term of the separation energy E as $\exp(-Et)$, we see that this energy effectively grows linearly with the distance r. Note that this is a perturbative result. Quarks will be trapped if the result remains true to all orders. The hope is that this holds in the strong coupling region and we shall prove in the next section that perturbation series converge in this region for sufficiently high dimensions and for some gauge groups.

On the other hand, this result could be false in the small coupling region (indeed, quantum electrodynamics does not confine electrons). For small enough coupling constant, one uses a steepest descent method around the maximum of $\chi(RRRR)/2g^2$, which arises from configurations where all $RRRR$ are equal to unity. The general solution of these equations depends on a arbitrary gauge transformation and reads

$$R_{xy} = \mathcal{R}_x^{-1}\mathcal{R}_y. \tag{3.2.1}$$

The fluctuations around these solutions are expected to be small in the small coupling region for non abelian gauge theories, at least at short distances, due to asymptotic freedom. The rotated matter fields $\Phi_x' = \mathcal{R}_x\Phi_x$ are then decoupled from the gauge fields in this approximation and now appear as free fields.

The difference between these two qualitative results seems to indicate that a phase transition might occur. The mean value $W(C)$ introduced in subsection 2.3 is a good tool for the study of this transition. It behaves as the exponentiated area in the "baryon phase" where quarks are trapped, and conversely as the exponentiated length of the loop in the "quark phase" where free quarks could be observed. The study of this transition is one of the main problems of the lattice theory. Of course, in dimension four, one expects the baryon phase to extend down to $g = 0$, which should therefore correspond to the transition point. We shall return to these problems in section 5.

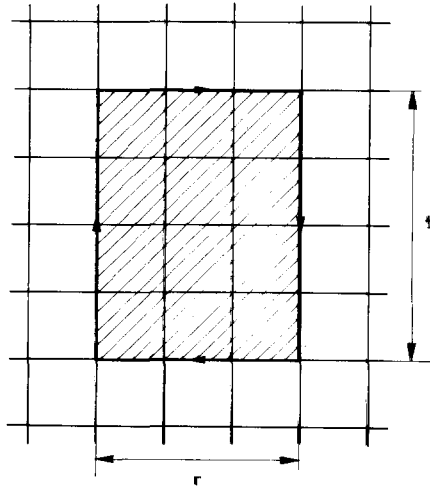There exists striking similitudes between the strong coupling expansion of lattice gauge theories and

Fig. 5. A diagram signalling confinement.

the relativistic string model [32]. In this phenomenological model, one studies the equations of motion of a string (a simple arc of curve), the energy of which is proportional to its length. Nambu [33] has introduced a Lagrangian formulation of this theory; the string sweeps a timelike surface and the action is proportional to the area of this surface. The gauge lattice model seems then to be a discretized version of the dual models. It adds nevertheless complementary terms for the motion of the ends of the strings and a natural interaction between strings (which appear rather artificially in the dual string model). The gauge lattice model could perhaps provide a bridge between the phenomenological dual models and conventional field theory.

In the same order of ideas, the Harari–Rosner duality diagrams [34] might be recovered. They would be considered as a sum over diagrams with some given (but not precised) topology for the surface, when one has given the quark line. The rules for writing such duality diagrams can be derived from the gauge lattice models. For instance, the Zweig–Okubo–Iizuka [35] empirical rule can be justified [36] in this scheme.

3.3. Diagrammatic techniques

In all our treatment, we always pass over in silence the problems related to the boundary conditions for the lattice. The gauge systems may allow the existence of several stable vacua, but we neglect here this fact. We always use a lattice of N nodes with periodic boundary conditions, and the fundamental state will not present topologically stable defects.

The graphs introduced in subsection 3.1 are drawn on the lattice. It is interesting to split the calculation in two steps. We introduce first the notion of *free graphs* or *diagrams*, with no labelling of nodes and links. A diagram is in fact the topological description of the complex formed by the graphs, without the support of an underlying lattice. All graphs pertaining to the same diagram yield the same factor, which is the contribution of the diagram. This is indeed independent of the lattice. The geometrical structure of the lattice, and in particular its dimension, only occurs in the number of graphs generated by a given diagram, i.e. the *number of configurations* of the diagram. For instance, in eq. (3.14), the cubic diagram gives a contribution $\Sigma_r(\beta_r/\beta_0)^6$ and can be embedded $N\, d(d-1)(d-2)/6$ times on a d-dimensional hypercubical lattice with N nodes. Rules for calculating numbers of configurations of 1-dimensional diagrams are well known in statistical mechanics [37]. These rules can be extended to higher dimensional cases [7] and are summarized in appendix B.

A further step is to restrict oneself to connected diagrams. Although this leads to important simplifications in usual statistical models, this is not as straightforward in the case of gauge theories. Indeed the rules for the calculation of connected diagrams entail contributions from many configurations which were not already present in the original expansion. For the numerical evaluation, the primitive expansion may in fact be more tractable. Nevertheless, the method is theoretically interesting. Indeed, the connected Green functions of the continuous theory are related to the connected correlation functions. Furthermore, this is the first step towards the construction of irreducible kernels which, first, lead to variational principles allowing the determination of critical points, and, secondly, permits the study of long distance behaviour with fewer terms in the expansion. Let us describe this procedure.

For our purpose, it is convenient to rewrite the strong coupling expansion in a slightly different manner. We restrict ourselves for clarity to pure gauge systems, but we introduce a source term K

and therefore consider the generating function

$$Z(\beta, K) = \int \left(\prod_{\langle xy \rangle} \mathscr{D}R_{xy}\right) \exp\left\{\beta \sum_p \chi(RRRR) + \sum_{\langle xy \rangle} \text{Tr}[^\text{T}K_{xy}D(R_{xy})]\right\}. \tag{3.3.1}$$

The character χ is the trace of the fundamental representation D and K_{xy} is a (complex) matrix in the corresponding representation space satisfying $K_{yx} = K_{xy}^\dagger$. For $\beta = 0$, Z factorizes in a product of unperturbed partition functions per link

$$z(K) \equiv \exp u(K) = \int \mathscr{D}R \exp 2 \text{Re} \text{Tr}[^\text{T}K D(R)]. \tag{3.3.2}$$

The exponentiated action is now expanded in powers of β and the resulting series is again interpreted in terms of graphs. A graph then is a finite set of lattice plaquettes with possible repetitions. Before the integration over the fields R, a plaquette $x\ y\ z\ t$ yields a term

$$P = \beta \text{Tr}[D(R_{xy})\ D(R_{yz})\ D(R_{zt})\ (R_{tx})]. \tag{3.3.3}$$

Of course, if it appears in the graph p times with this orientation and q times with the reversed one, this factor will be $P^p P^{*q}/(p!q!)$ where the denominator is just a symmetry number.

Each factor $D(R)^{\alpha\beta}$ corresponding to a given link can be obtained by a derivative $\partial/\partial K^{\alpha\beta}$ acting on the right-hand side of eq. (3.3.2). Hence an oriented link shared by $p + q$ plaquettes (p with the right orientation, q with the reverse) contributes a term

$$\langle D(R)^{\alpha_1\beta_1} \ldots D(R)^{\alpha_p\beta_p} D(R^{-1})^{\gamma_1\delta_1} \ldots D(R^{-1})^{\gamma_q\delta_q}\rangle_0$$

$$= z(K)^{-1} \frac{\partial}{\partial K^{\alpha_1\beta_1}} \cdots \frac{\partial}{\partial K^{\alpha_p\beta_p}} \frac{\partial}{\partial K^{\gamma_1\delta_1*}} \cdots \frac{\partial}{\partial K^{\gamma_q\delta_q*}} z(K), \tag{3.3.4}$$

or, in a shorthand notation which saves a lot of writing,

$$\langle R^{(p,q)}\rangle_0 = z(K)^{-1} \frac{\partial^{p+q}}{\partial K^p \partial \bar{K}^q} z(K). \tag{3.3.5}$$

The final value of the graph is given by the product of link contributions (3.3.5) and plaquette contributions β, divided by an overall factor equal to the order of the symmetry group of plaquette interchange. As previously, one may compute the contribution of each diagram (free graph not drawn on a lattice) and multiply by the number of configurations of this diagram. The sum over all possible distinct diagrams (including the empty one, contributing a term equal to unity) reproduces the expansion of Z/Z_0 (with $Z_0 \equiv Z(0, K) = \prod_{xy} z(K_{xy})$). We give in appendix C the functions $z(K)$ for some classical groups and an example of the Z expansion.

The preceding expansion of subsection 3.1 is retrieved by summing over all the number of plaquettes at a given location, and rearranging the result following the irreducible representations of the gauge group. This is then considered as a single plaquette carrying a representation index. In this picture $z(K)$ is a generating function for the Clebsch–Gordan series.

Next we turn to connected diagrams, which are related to the logarithm F of the partition function Z. In the preceding embedding of diagrams on the lattice, it is forbidden to map two different links of the diagram at the same location. It is necessary to circumvent this interdiction in order to restrict ourself to connected diagrams. For this purpose, one introduces new factors for a link. Plaquettes sharing this link are tied together in all possible ways and we define *connected average* $\langle R^{(p,q)}\rangle_c$ by the standard definition
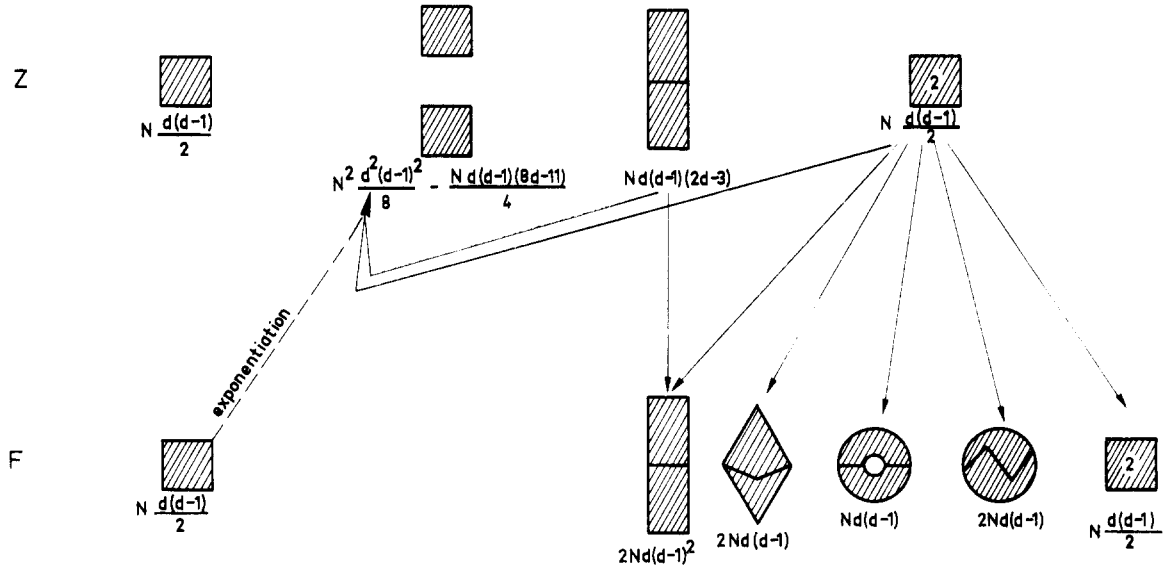
Fig. 6. The introduction of connected diagrams.

$$\langle R^{(p,\,q)}\rangle_0 = \sum \langle R^{(p_1,\,q_1)}\rangle_c \ldots \langle R^{(p_k,\,q_k)}\rangle_c,$$ (3.3.6)

where the summation extends over all partitions of p and q. It is easy to verify that the explicit result is

$$\langle R^{(p,\,q)}\rangle_c = \frac{\partial^{p+q}}{\partial K^p \partial \bar{K}^q} u(K).$$

If one uses this contribution rather than (3.3.5) for a link, and if the diagrams are now allowed to map several links onto the same lattice location (caution: the rules of appendix B are no longer valid), summations over all connected parts can now be performed independently. As a result, the logarithm F of the partition function may be expressed entirely in terms of connected diagrams

$$F \equiv \log Z = \sum_{\text{links}} u(K) + \sum (\text{connected diagrams}).$$ (3.3.7)

Figure 6 illustrates the procedure up to second order diagrams. Arrows indicate the action of eq. (3.3.6). It is noteworthy that new diagrams now appear. The number of configurations, which differs in the two calculations for a given diagram, is also indicated.

It is possible to go beyond the connected expansion by considering irreducible diagrams [38]. A connected diagram is irreducible if it cannot be separated by a cut along a link. In fig. 6, this is the case for all connected diagrams in the expansion of F, except for the first diagram of second order. A general connected diagram is thus a tree of irreducible parts (fig. 7). Suppose there are n_B irreducible parts connected by n_L links and let n_M be equal to the number of these links, each one counted as many times as the number of irreducible parts it connects. The following topological identity [39]

$$1 = n_B + n_L - n_M$$

may be averaged as follows. Let $B(\{M\})$ be the sum of contributions over irreducible diagrams computed according to the rules for connected diagrams, except for the replacement of $\langle R^{(p,\,q)}\rangle_c$ by the
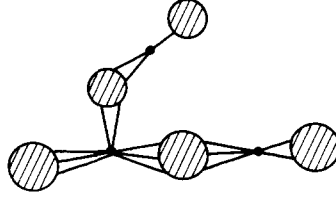
Fig. 7. A connected diagram seen as a tree of irreducible parts. Lines converging to a point represent plaquettes sharing a link.

variables $M^{(p,q)}$. If these variables satisfy the system of equations

$$G_{xy}^{(p,q)} = \frac{\partial B}{\partial M_{xy}^{(p,q)}}(\{M\}),\tag{3.3.8}$$

$$M_{xy}^{(p,q)} = \left[\exp\left(\sum_{k+l\geqslant 1} G_{xy}^{(k,l)}\frac{\partial^{k+l}}{\partial K^k \partial \bar{K}^l}\right)\right]\langle R_{xy}^{(p,q)}\rangle_c\tag{3.3.9}$$

which may be graphically interpreted as shown in fig. 8, $B(\{M\})$ will just reconstruct the tree of fig. 7, counted n_B times. Thus the average of eq. (3.3.7) reads

$$F = B(\{M\}) + \sum_{\langle xy\rangle} M_{xy}^{(0,0)} - \sum_{\langle xy\rangle}\sum_{p+q\geqslant 1} M_{xy}^{(p,q)} G_{xy}^{(q,p)}.\tag{3.3.10}$$

The set of eqs. (3.3.8), (3.3.9) and (3.3.10) characterizes indeed a generalized Legendre transformation. The conjugate variables are $M_{xy}^{(p,q)}$ and $\langle R_{xy}^{(p,q)}\rangle_c$ for $p + q \geqslant 1$. B depends only on the first set, and F on the second set of these variables. It is easy to verify that eqs. (3.3.8) and (3.3.9) indicate that F, given in (3.3.10) is stationary with respect to the variations of $\{M\}$ and $\{G\}$. As is well known, such a Legendre transformation can be inverted and B may be calculated from (3.3.10) by writing the stationarity conditions with respect to the variations of $\{G\}$ (using again eq. (3.3.9)) and $\{\langle R_{xy}^{(p,q)}\rangle\}$:

$$\sum_{j}\sum_{\substack{k_1+\cdots+k_j=p\\ l_1+\cdots+l_j=q\\ k_i+l_i\geqslant 1}}\frac{G_{xy}^{(k_1,l_1)}\cdots G_{xy}^{(k_j,l_j)}}{j!} = \frac{\partial F}{\partial\langle R_{xy}^{(p,q)}\rangle_c}.$$

Of course, in all these considerations, $M^{(0,0)}$ and $\langle R^{(0,0)}\rangle$ must not be considered as free variables.

It should be possible to take advantage of this formalism for the determination of Green functions. The reconstruction of trees of the type indicated on fig. 7 is a resummation technique, and allows the propagation of states at long distances, even at low orders in the perturbative calculation. This was not the case with the primitive expansion since the kinetic term is treated only perturbatively. And

$$G_{xy}^{(n)} =$$

n plaquettes
sharing the link x y

$$M_{xy}^{(n)} =$$

n plaquettes
leaving the link x y

Fig. 8. Graphical interpretation of eqs. (3.3.8) and (3.3.9) (for clarity, orientations are omitted).

indeed the perturbative series for free actions, as given in eq. (2.3.2) can be exactly summed by this technique (in this case, there is only one irreducible diagram). We do not develop this possibility in more detail and apply it only for the determination of critical points and phase transitions of our models.

3.4. Determination of critical points

We have seen in the preceding paragraph that F is the stationary value of some function Φ ($\{x_i\}$) (the right-hand side of eq. (3.3.10)) when the parameters x_i (here the M's and G's) are varied. This variational principle allows the determination of critical points. Suppose there are two distinct stationary points. The corresponding values for F depends on the coupling constant β and may cross-over for some critical value β_c (see fig. 9a). At this point, the parameters, and thus the properties of the system, have a discontinuity. This fact characterizes a first-order transition. The determination of such a critical behaviour follows from the system:

$$\begin{cases} \Phi(\{x_i^c\}) = \Phi(\{y_i^c\}) \\ \dfrac{\partial \Phi}{\partial x_i}(\{x_i^c\}) = \dfrac{\partial \Phi}{\partial y_i}(\{y_i^c\}) = 0 \end{cases} \tag{3.4.1}$$

which allows the computation of β_c and of the critical parameters (x_i jumping from x_i^c to y_i^c).

Another possibility is the appearance of new extrema for some critical value (see fig. 9b). The variational parameters then depart smoothly from their original values (with only a break in the slope) and this is characteristic of a second-order transition. The conditions read

$$\begin{cases} \dfrac{\partial \Phi}{\partial x_i}(\{x_i^c\}) = 0 \\ \det\left[\dfrac{\partial^2 \Phi}{\partial x_i \partial x_j}(\{x_i^c\})\right] = 0. \end{cases} \tag{3.4.2}$$

Unfortunately, Φ is only obtained as a series in β. The determination of critical values will then be done in two steps. First, one looks for the nature and the position β_c^0 of the transition at lowest order. Then the critical quantities are computed perturbatively as series in β_c^0 at a given order. Such a
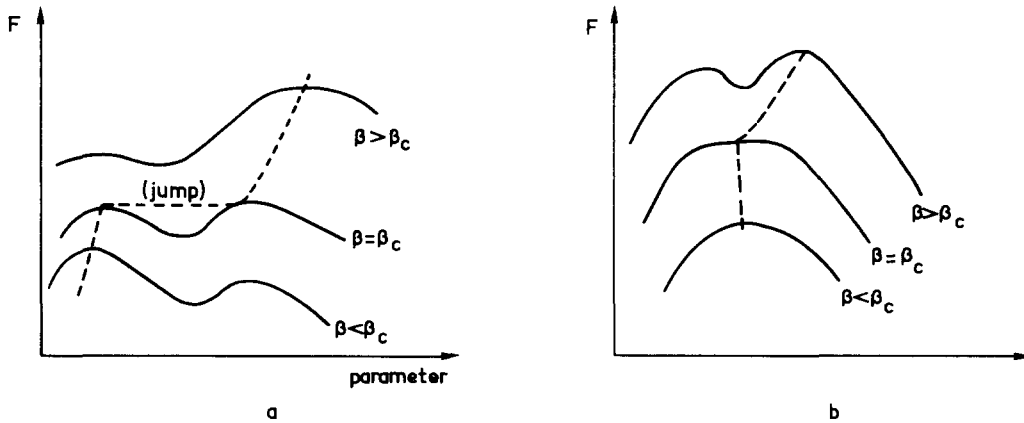

Fig. 9. The characterization of a first order (a) or a second order (b) transition.

procedure can of course be misleading and has to be checked whenever possible using limiting situations.

It must be emphasized that much care must be taken in the choice of the variational parameters. A transition may indeed escape the investigation if the set of parameters is too restricted. It is also possible that the perturbative determination has no meaning for lack of convergence. We do not give here a general discussion of these problems and, in particular, postpone up to the following sections the discussion of the choice of an order parameter. Nevertheless, these questions will reappear in the results of the example now treated. We consider the simple case of a Z_2 gauge group in which the number of free variables has been (naively?) reduced by equating M's and G's values on different links. This calculation appears therefore as a mean field analysis, with the usual drawback of such a procedure. It should however be a sensible method in very high dimension and we take it only as a guide towards a more refined treatment. In the Z_2 case, the notations are simplified since the representations are real (therefore plaquettes need no orientation) and one dimensional. The matrix K reduces to a real number and the free energy per link is $u(K) = \text{Ln cosh } K$ (see appendix C).

At lowest order, $\mathscr{B}$ reduces to a sum over single plaquette graphs:

$$\mathscr{B} = \beta \sum_p M^{(1)} M^{(1)} M^{(1)} M^{(1)} + O(\beta^2). \tag{3.4.3}$$

Thus only the variables $M_{xy}^{(1)}$ and $G_{xy}^{(1)}$ appear. The differential operator used in (3.3.9) reduces indeed to a simple translation of $G^{(1)}$ of an amount equal to K. Equating now all $M_{xy}^{(1)}$ to M, and all $G_{xy}^{(1)}$ to G, the system is written

$$F = \mathop{\text{Extr}}_{M,G} \left\{ \frac{Nd(d-1)}{2} \beta M^4 + Ndu(K+G) - NdMG \right\}. \tag{3.4.4}$$

The extremum condition over M

$$G = 2\beta(d-1)M^3$$

allows to eliminate G and we finally get, in the absence of external sources ($K = 0$):

$$\frac{F}{Nd} = \mathop{\text{Extr}}_{M} \{ u(\beta^* M^3) - \tfrac{3}{4}\beta^* M^4 \} \quad \text{with} \quad \beta^* = 2(d-1)\beta. \tag{3.4.5}$$

We have displayed the term between braces as a function of M for different values of β^* in fig. 10. The behaviour is characteristic of a first-order transition which occurs for $\beta_c^{0*} = 2.7552$. This situation is in fact present whatever the gauge group may be. The eq. (3.4.5) remains indeed true, with different functions u. One cannot leave the extremum at $M = 0$ smoothly since the two terms in (3.4.5) are of different order near $M = 0$ and thus cannot compete for reversing the curvature.

At the next order, we use all the irreducible diagrams of fig. 6. We observe now the appearance of $M^{(2)}$ and $G^{(2)}$ variables, which are of order β^2. The corrections on the critical values can be computed easily, but we omit here the rather long algebra and only present the results [7].

In table 1, the critical couplings are displayed for some gauge groups (Z_2, U(1) and SU(2)) and some dimensions. In a few cases, the values can be compared with the known results. The agreement seems to be excellent in high dimension. Let us recall that, in statistical models, the perturbation expansion can be rearranged as a series in d^{-1}[40]. For the gauge systems, this is not true and it can be verified that each power of d^{-1} involves an infinite number of diagrams, except the lowest one (mean field result).
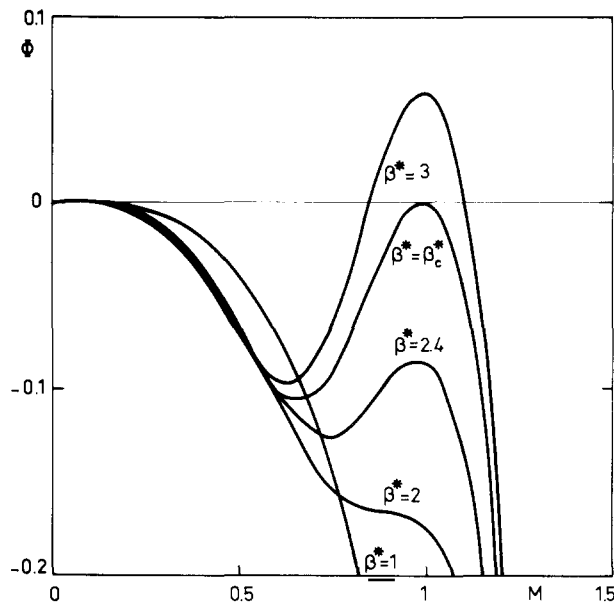
Fig. 10. The first order transition in Z_2 pure gauge system.

Some discrepancy is present at $d = 3$ for the Z_2 system. In this case however the transition is of second order [21], while the present calculation is performed in the framework of a first-order one. We expect indeed that such a calculation is at the limit of convergence and credibility. This can be checked by treating the case $d = 2$, which is equivalent to a one-dimensional Ising model and therefore does not present any transition. Whereas the lowest order predicts a spurious transition, the first correction is sufficient to rule out this phenomenon.

Table 1

group	approx.	$d = 3$	$d = 4$	$d = 5$	$d = 6$	$d = \infty$
Z_2	0	2.6028	2.6840	2.7104	2.7229	2.7552
	2	4.2970	3.5272	3.2905	3.1677	
	3	4.2941	3.5257	3.2897	3.1672	
	exact	4.5678	3.5254	?	?	
U(1)	0	5.6563	6.2133	6.4793	6.6384	7.2934
	2		11.3761	9.0530	8.4929	
	Padé	10.0934	8.9781	8.4968	8.2298	
SU(2)	0	5.8708	6.7270	7.1520	7.4096	8.4787
	2			12.821	10.5948	
	Padé		11.4160	10.333	9.8510	

The conclusions of our analysis are the following. In any case, there seems to appear a first-order transition in the pure lattice gauge system for a high enough dimension. However there exists a lower critical dimension d_c for which this transition disappears. d_c might differ for abelian and non-abelian systems; in the non-abelian case, d_c might be four as suggested by asymptotic freedom. One would always be in the confined phase and there would be hope to understand (at least qualitatively) the properties of the continuous system by studying the lattice strong coupling expansions.

3.5. Mass spectrum

We have constructed in subsection 2.4 a space for the unperturbed states of the theory. Let us now describe the perturbative determination of the energy levels. We use the Hamiltonian formulation and restrict ourselves for clarity to a pure gauge field. The discussion is carried in three spatial dimensions for a SU(3) color gauge group [8]. The Hamiltonian (2.4.7) is split into an unperturbed part

$$H_0 = \frac{g^2}{a^{d-3}} \sum_{\langle xy \rangle} \frac{l_{xy}^{i2}}{\chi(l_{xy}^{i2})} \tag{3.5.1}$$

and a perturbation vanishing for strong couplings

$$V = \frac{a^{d-5}}{2g^2} \sum_p \chi(RRRR). \tag{3.5.2}$$

H_0 is diagonal in the space built with string bits states, and its unperturbed spectrum has already been displayed (fig. 4). We have first to take care of degeneracies. For instance, the lowest unperturbed state (the boxciton) is $6N$ times degererate. However this degeneracy is easily analysed in terms of the translation-rotation group of the lattice. We first combine the different orientations as shown in fig. 11. The resulting states $|A_1\rangle$, $|F_1\rangle$ and $|E\rangle$ belong respectively to a one-, three-, and

Fig. 11. Combining different orientations of the boxciton.

two-dimensional irreducible representations of the point group O (rotation group of the cube: we use the cristallographic Schoenflies notations). The translation invariance is now taken into account by summing on the different locations x with a weight factor exp (i $k \cdot x$). The basis $|A_1, k\rangle$, $|F_1, k\rangle$ and $|E, k\rangle$ is now well adapted to the perturbative treatment. Let us mention that the decomposition of SO(3) representations in irreducible representations of the O group indicates that spin 0, 1 and 2 states are contained only in A_1, F_1 and E levels respectively.

In fact, the knowledge of an energy level $\mathscr{E}(k)$ allows the introduction of two different "masses". One is the static energy $\mathscr{E}(0)$, the second is the curvature of the level at the point $k = 0$. The two masses should become identical in the limit of a zero lattice spacing where the Lorentz symmetry is expected to be retrieved. The unperturbed Hamiltonian H_0 has no kinetic term and therefore does not give any curvature to $\mathscr{E}(k)$. The perturbation V is responsible for propagation and therefore many terms are needed in order to build this curvature mass. This is no surprise since the strong coupling expansion treats kinetic energy as a perturbation. It is easier to compute $\mathscr{E}(0)$, which is obtained by summing over all possible locations of the final state.

After these preliminary remarks, we now carry the perturbative calculation of the energy levels. They are obtained by the diagonalization (in the space of all the unperturbed states of energy E_0) of the operator [41]

$$H_0 + V + V\frac{1-P_0}{E_0-H_0}V + \left\{V\frac{1-P_0}{E_0-H_0}V\frac{1-P_0}{E_0-H_0}V - V\frac{1-P_0}{(E_0-H_0)^2}VP_0V\right\} + \dots \tag{3.5.3}$$
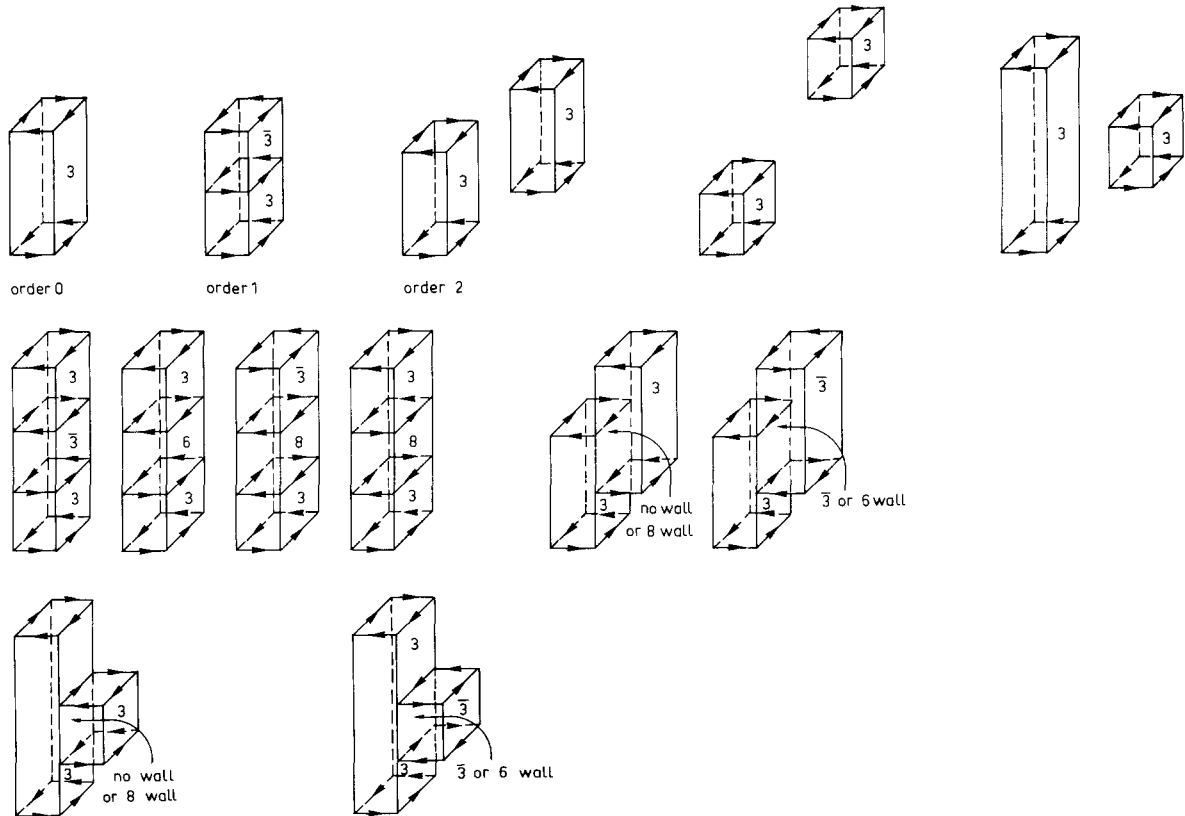

Fig. 12. Graphs contributing up to second order to the energy level of the boxciton.

where P_0 is the projector over the unperturbed levels. This operator is already diagonal (by blocks) in the basis A_1, F_1, E. It is possible to represent graphically eq. (3.5.3). Each V is represented by an oriented square at some time, $(1 - P_0)/(E_0 - H_0)$ or P_0 the surface swept by the string bit states during time. In fig. 12 are given the graphs contributing up to second order to the energy of the boxciton. Gathering all these contributions, one gets [8]

$$\begin{cases} \mathscr{E}_{A_1} = \tfrac{16}{3}(1 - y - 0.5685\, y^2 + \cdots) \\ \mathscr{E}_{F_1} = \tfrac{16}{3}(1 + y + 0.0971\, y^2 + \cdots) \\ \mathscr{E}_{E} = \tfrac{16}{3}(1 - y + 0.582\, y^2 + \cdots) \\ \text{with } y = 3/8g^4 \,. \end{cases} \tag{3.5.4}$$

With matter fields, a similar analysis can be performed. We do not give here the detailed calculations which can be found in original papers [42,43,8,44]. We only report and discuss their main results.

Let us turn first to Schwinger's model. The exact results in the continuous theory are a severe test for the perturbative lattice calculations. The Cornell group [42, 43] has pushed up to eighth order and has obtained the following striking results. With the notations

$$x = 1/g^2 a^2, \qquad \mu = (2m/g)\sqrt{x}, \qquad \alpha = 1 + 2\mu, \qquad \beta = 3 + 2\mu, \qquad \gamma = 1 + \mu$$

they get for the vacuum energy, vector and scalar bound state masses respectively:

$$\frac{E_0}{N} = -\frac{1}{\alpha}x^2 + \frac{3}{\alpha^3}x^4 - \frac{58 + 40\mu}{\alpha^5 \beta}x^6 + \frac{(5772 + 13802\mu + 10940\mu^2 + 3000\mu^3 + 80\mu^4)}{4\alpha^7 \beta^2 \gamma}x^8 + \cdots$$

$$\frac{2\sqrt{x}M_V}{g} = 1 + 2\mu + \frac{2}{\alpha}x^2 - \frac{(10 + 4\mu)}{\alpha^3}x^4 + \frac{(236 + 272\mu + 96\mu^2 + 16\mu^3)}{\alpha^5 \beta}x^6$$
$$- \frac{(6626 + 18738\mu + 19980\mu^2 + 10244\mu^3 + 2808\mu^4 + 464\mu^5 + 32\mu^6)}{\alpha^7 \beta^2 \gamma}x^8 + \cdots$$

$$\frac{2\sqrt{x}M_s}{g} = 1 + 2\mu + \frac{6}{\alpha}x^2 - \frac{(26 + 4\mu)}{\alpha^3}x^4 + \frac{(572 + 448\mu + 16\mu^2 - 16\mu^3)}{\alpha^5 \beta}x^6$$
$$- \frac{(15810 + 32188\mu + 32284\mu^2 + 8964\mu^3 - 262\mu^4 - 176\mu^5 + 32\mu^6)}{\alpha^7 \beta^2 \gamma}x^8 + \cdots.$$

From these expansions, one gets:

a) for $g = 0$ at fixed a^2, the ground state of a $x - y$ antiferromagnetic system. The [2/2] Padé approximants allow to study the infinite x limit:

$$E_0/N = -0.665([2/2]\ \text{Padé}), \text{ exact result: } -0.637.$$

b) In the massless Schwinger model (obtained for $m = 0$, $x \to \infty$), the scalar particle decouples from the theory, and M_s/M_V must be equal to 2. The [2/2] Padé approximant for this quantity gives 1.95 which lies within 2.5% of the exact answer.

c) The vector particle has a mass $g/\sqrt{\pi}$. Forming the [2/1] Padé approximant of the fourth power of $2\sqrt{x}M_V/g$ expansion, one obtains:

$$M_V/g = 0.769, \text{ exact result: } 0.546.$$

Clearly, this is less accurate (36%), but represents a considerable improvement over a lower order

calculation (which gives 0.841). It indicates that it is more difficult to obtain accurate masses rather than mass ratios.

d) The ratio between the two different masses (curvature of the energy level over zero momentum energy) has also been computed and is found to be equal to 0.86 at this order.

All these results are particularly encouraging. It seems nevertheless that relatively high order (about 8, or more) are necessary for obtaining sensible values. This needs an important computer machinery which should be very useful for more realistic cases (Wilson's model). The first step in this direction is the study of a pure SU(3) gauge field. We have outlined in this paragraph the calculations which have been pursued at the present time up to 4th order [8].

4. Global properties and critical points

4.1. Mean field analysis

In subsection 3.4, we have developed a perturbative variational method used in particular in the determination of critical points. The variables M and G introduced in this context act as some mean fields and the irreducible function $\mathscr{B}$ is an effective action with these mean fields. We retrace in this paragraph the lowest order approximation, in a more general framework as the one of the perturbative expansion.

The mean field approximation is a standard method in statistical mechanics. Assume that the action is linear in the field Φ_i located at site i of some lattice, i.e. reads

$$S(\Phi_i) = \sum_i \Phi_i K_i \tag{4.1.1}$$

where K_i depends of course on the fields located at other sites. In this way of writing, K_i behaves as some external field and is approximated by some constant mean field K (to be computed consistently). This is expected to give good results if many terms enter the definition of K_i, in order to smear the fluctuations, that is, in our models, if the dimension d of the lattice (or more generally its coordination number) is large enough.

To give some ground to this expectation, one can proceed as follows. The convexity of the exponential function allows to write the Peierls inequality [45]:

$$\left\langle \exp\left(S(\Phi) - K\sum_i \phi_i\right)\right\rangle \geq \exp\left(\left\langle S(\Phi) - K\sum_i \Phi_i\right\rangle\right). \tag{4.1.2}$$

In particular, one expects to be near equality if the approximation described above is valid. The measure defining the mean values $\langle X\rangle$ in (4.1.2) is left arbitrary, we choose

$$\langle X\rangle = \exp(-Nu(K))\int \prod_i \mathscr{D}\Phi_i \exp\left(K\sum_i \Phi_i\right), \tag{4.1.3}$$

with

$$\exp(u(K)) = \int \mathscr{D}\Phi \exp K\Phi$$

in order to retrieve on the left-hand side of the inequality the partition function $Z = \exp F$. The free energy F thus obeys an inequality

$$F \geq Nu(K) + \langle S(\Phi)\rangle - NK\langle \Phi\rangle. \tag{4.1.4}$$

The mean value $\langle\Phi\rangle$ is nothing but $u'(K) = \mathrm{d}u(K)/\mathrm{d}K$. Since the inequality is valid for any K and since F does not depend on K, we can maximize the right-hand side with respect to K. Thus one finally gets

$$F \geq \underset{K}{\mathrm{Max}}\{Nu(K) + S(u'(K)) - NKu'(K)\}. \tag{4.1.5}$$

The right-hand side is what we call the mean field approximation. Irrespective of symmetries or other detailed properties of the model, it always gives a lower bound for F. Note that the extremum property obtained in subsection 3.4 is now made precise as a maximum property.

As the first order transition of a pure gauge field has already been discussed at length, we turn to a more complex case involving other fields. We consider the spontaneous symmetry breaking initiated by the action (2.5.7). Two mean fields are introduced: one, denoted by K, for the gauge field, the other, H, for Higgs field. In this case, K is a matrix in the fundamental representation space of the gauge group. The function $u(K)$ have already been introduced and are tabulated for some usual groups in appendix C. The corresponding function for the matter fields are written $v(H)$ and one gets

$$\frac{F}{N} \geq \underset{K,H}{\mathrm{Max}} \left\{ \alpha dv'(H)u'(K)v'(H) + \beta\frac{d(d-1)}{2}\mathrm{Tr}[u'(K)^4] + v(H) - Hv'(H) + d[u(K) - Ku'(K)] \right\}. \tag{4.1.6}$$

The study [6] of such a formula shows the existence of three regions in the (H,K) plane. A strong coupling region does not present any spontaneous symmetry breaking ($H = K = 0$) and is separated from the other phases by a first order transition line. For low g values, the upper maximum arises for non zero K. As α reaches some critical value, the maximum becomes a saddle point. There exists thus a second order transition line separating the unconfined phase in massive ($H \neq 0$, $K \neq 0$) and non-massive ($H = 0$, $K \neq 0$) regions. Figure 13 illustrates the phase diagram obtained in the case of a $U(1) \sim SO(2)$ gauge group (drawn for large d when one expects the validity of the mean field approximation).
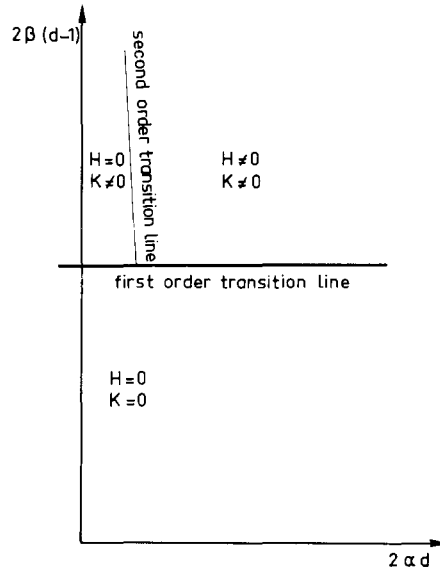

Fig. 13. Phase diagram for the U(1) gauge system (d large).

One may question the validity of the qualitative results obtained through mean field theory. Indeed the mean field is the conjugate variable to some order parameter. The analysis is expected to be right only if a realistic order parameter has indeed been identified. Here a local order has been encountered. In fact, the realistic order is probably to be described in terms of global quantities such as $W(C)$ introduced in subsection 2.3. At the present time, the use of such order parameters in this context has not been seriously attempted.

Furthermore, the vacuum is highly degenerate due to local gauge invariance. A doubt therefore remains on the validity of the method since fluctuations would spoil any attempt to improve the approximation by some perturbative expansion.

4.2. Exact two-dimensional solutions

After studying high dimensionalities, we turn now to low dimensions where some exact results can be found. They should enable one to test the validity of the general features encountered in the previous subsections. Let us consider a two-dimensional lattice; we do not assume periodic boundary conditions and take a square of $N = L^2$ lattice sites. Within this framework, the plaquette variables $P = RRRR$ are independent. To see this, one might start from the edges of the square and integrate successively on the free variables pertaining to the boundary links. This leads to

$$Z = \left\{ \int \mathscr{D}R \, \exp[\beta\chi(R)] \right\} = \exp\left\{ Nu(\beta) \right\}. \tag{4.2.1}$$

Hence the partition function is exactly known. It is free from any singularity and there is no critical behaviour in this two-dimensional pure gauge model for finite couplings. This result is a generalization of the well known solution for one-dimensional Ising (and more generally classical Heisenberg) model.

It is possible to compute exactly for such two-dimensional models any mean value. For instance, let us consider the average $W(C)$ introduced in subsection 2.3. The method of subsection 3.1 is directly applicable in this case. Only one graph satisfies the selection rules (namely the one represented in fig. 5) and hence

$$W_r(C) = \left\langle \chi^r \left(\prod_{\langle xy \rangle \in C} R_{xy} \right) \right\rangle = \beta_r^{s(C)} \tag{4.2.2}$$

where $s(C)$ is the surface (i.e. the number of plaquettes) enclosed by the curve C. The average value decreases exponentially with this area. This is what we expected in general in the disordered phase. Note that if one uses in (4.22) the fundamental representation, β_r is easily computed in terms of the function u and we finally get in this case:

$$W(C) = \exp\left\{ -s(C) \, \mathrm{Log} \frac{1}{\mathrm{d}u(\beta)/\mathrm{d}\beta} \right\}. \tag{4.2.3}$$

We note that, as β goes to infinity (i.e. $g \to 0$), the coefficient of the exponential decrease vanishes as can be verified for the usual gauge groups. The quark binding then becomes less and less effective when one approaches the point $\beta = \infty$ which acts as a pseudo-critical point.

4.3. The gauge group Z_2

The physical interest of the lattice theory lies in its continuous limit where one recovers the Yang–Mills field theory if one uses a continuous Lie group as gauge group. The study of the Z_2

system, which is without interest in this point of view, is nevertheless very instructive. Some properties of the lattice systems are in fact common to discrete and continuous groups. Furthermore, several devices introduced in the context of the Ising model can be used in Z_2 gauge systems [21]. The first is the duality transformation which allows exact predictions for dimension 3 and 4. Furthermore, it is possible to give an almost complete proof of the existence of a transition with $W(C)$ acting as order parameter.

Geometrical duality [46] transforms p-dimensional manifolds into $d-p$ dimensional ones. A lattice dual to the original one is introduced through a translation by an amount $(\frac{1}{2}a, \frac{1}{2}a, \dots)$. To a site is associated the dual hypercube surrounding it. To a link corresponds the $(d-1)$-dimensional face belonging to both hypercubes associated to the endpoints, and so on.

We now apply the transformation to a Z_2-gauge system with sources. The partition function reads

$$Z = 2^{-Nd} \sum_{\{R_{xy}=\pm 1\}} \exp\left\{\alpha \sum_{xy} R_{xy} + \beta \sum_p RRRR\right\}$$

$$= 2^{-Nd}(\cosh \alpha)^N (\cosh \beta)^{Nd} \sum_{\{R_{xy}=\pm 1\}} \prod_{xy} (1 + \tanh \alpha R_{xy}) \prod_p (1 + \tanh \beta RRRR) \qquad (4.3.1)$$

and we recover the diagrammatic expansion of subsection 3.1. Non-vanishing terms are in one-to-one correspondence with p distinct plaquettes of the lattice. The boundary is the set of L links which belong to an odd number of selected plaquettes. A configuration contributes a term $(\tanh\beta)^p (\tanh\beta)^L$.

We introduce dual fields S associated to $(d-2)$-dimensional elementary cells of the dual lattice. Then each S corresponds to a given plaquette. Given a graph contributing to Z, let $S = -1$ if the corresponding plaquette is selected, $S = 1$ otherwise. A one-to-one correspondence between strong coupling expansion graphs and configurations on the dual lattice has indeed been realized. Obviously $P = \frac{1}{2}\Sigma_i(1 - S_i)$. As a link pertains to $2(d-1)$ plaquettes, or, in the dual language, a $(d-1)$-dimensional hypercube c is bounded by $2(d-1)$ faces, we also have $L = \frac{1}{2}\Sigma_c(1 - \Pi_{s \in \partial c} S)$. Finally, the partition function may be rewritten in terms of dual configurations

$$Z = 2^{-Nd}(\cosh\alpha)^N(\cosh\beta)^{Nd} \sum_{\{S=\pm 1\}} \exp\left\{(\tfrac{1}{2}\text{Ln}\tanh\alpha)\sum_c \left(1 - \prod_{s \in \partial c} S\right) + \tfrac{1}{2}\text{Ln}\tanh\beta\sum_i(1 - S_i)\right\}. \qquad (4.3.2)$$

Let us now use this construction in some specific dimensions:

i) $d = 2$. The dual fields S are located on the nodes of the dual lattice and the link c involves two neighbouring sites. Hence (4.3.2) is the partition function of the usual two-dimensional Ising model with an external field H:

$$Z_{\text{Ising}}(\beta^1, H) = 2^{-N} \sum_{\{S_i=\pm 1\}} \exp\left\{\beta^1 \sum_{\langle ij\rangle} S_i S_j + H \sum_i S_i\right\} \qquad (4.3.3)$$

and we get the equality

$$F(\alpha,\beta) = \tfrac{1}{2}\text{Ln}(\sinh^2 2\alpha \cdot \sinh\beta) + F_{\text{Ising}}(-\tfrac{1}{2}\text{Ln}\tanh\alpha, -\tfrac{1}{2}\text{Ln}\tanh\beta). \qquad (4.3.4)$$

The two-dimensional Ising model is well known. It exhibits no transition for $H \neq 0$, a second-order transition at $H = 0$, $\beta_c^1 = \text{Ln}(1 + \sqrt{2})$. This is in agreement with the results of the preceding subsection: no transition occurs for finite α, β. The point $\alpha = 0$, $\beta = \infty$ acts as a second-order critical point.

ii) $d = 3$. The dual fields are now located on links. The model is self dual, and, more precisely

$$F(\alpha,\beta) = \tfrac{3}{2}\text{Ln}(\sinh 2\alpha \cdot \sinh 2\beta) + F(-\tfrac{1}{2}\text{Ln}\tanh\beta, -\tfrac{1}{2}\text{Ln}\tanh\alpha). \qquad (4.3.5)$$

Note the interchange of the roles of α and β in the two sides of the formula. The phase diagram of the system, and, in particular the first-order and second-order transition lines must obey this symmetry. For instance, the pure gauge system ($\alpha = 0$) is dual to the Ising model ($\beta = \infty$): this infinite value forces the product of fields bordering any plaquette to be equal to unity. It has been already noticed that the general solution is $S_{xy} = \sigma_x \sigma_y$ and this replacement leads again to the Ising action. As a consequence, it presents a *second-order* transition for $-\frac{1}{2}$ Ln tanh $\beta_c = \beta_c^{1,d=3}$ (i.e. $\beta_c \sim 0.7613$, using numerical results on the Ising model in dimension three). It is therefore proved that the Z_2 gauge system, which has no transition for $d = 2$, presents for $d = 3$ a second-order transition (Recall the former analysis: a first order transition is expected for high dimensions).

iii) $d = 4$. The dual model is of a new type. Dual fields S are located on plaquettes and the interaction involves the six plaquettes bounding each three-dimensional cube. As α goes to zero, this interaction becomes infinite: then the product of the six plaquette fields is constrained to be unity. Such constraints can be solved by cohomology methods and one finally finds that the S_{ijkl} field located on the plaquette $ijkl$ is of the form $S_{ijkl} = R_{ij}^* R_{ij}^* R_{kl}^* R_{li}^*$ where new fields $R_{ij}^* = \pm 1$ located on links ij have been introduced. Using this expression, we recover the original pure gauge field model. The final result is expressed as a self duality formula

$$F(0,\beta) = 3\text{Ln sinh}2\beta + F(0, -\tfrac{1}{2}\text{Ln tanh}\beta). \tag{4.3.6}$$

This remarkable result is analogous to the Kramers–Wannier duality for the two-dimensional Ising model. If we assume that the system undergoes a unique transition (as for $d = 3$), then the critical constant follows

$$\beta_c = -\tfrac{1}{2}\text{Ln tanh }\beta_c, \quad \text{or} \quad \text{sinh }2\beta_c = 1. \tag{4.3.7}$$

It is in fact possible to give a proof of the existence of a transition in the Z_2 gauge system by studying the behaviour of Log $W(C)$ for large curves C. More precisely [21], let $s(C)$ denote the minimal area enclosed by C. If $d \geqslant 3$ and for sufficiently low β, one can prove that there exist two positive constants, a_1 and a_2 such that

$$a_1 \leqslant -\frac{\text{Ln } W(C)}{s(C)} \leqslant a_2.$$

The inequality on the right-hand side is obtained by finding an upper bound for the terms of the strong coupling series for $W(C)$. The result holds beyond the perturbative framework since one finds a convergent series with positive terms. The inequality on the left-hand side is a consequence of the Griffiths–Kelly–Sherman result [47].

In order to prove the existence of the transition, it would be nice to show that, for β large enough and $d \geqslant 3$, $-$Ln $W(C)$ behaves as the length of C. In fact, it is sufficient to establish it for $d = 3$, since an argument based again on the Griffiths–Kelly–Sherman inequalities extends this result to higher dimensions. The result is true perturbatively (as is shown by the strong coupling expansion for the dual function of $W(C)$). Unfortunately no complete proof beyond the perturbation theory exists at the present time.

5. Recurrence relations

5.1. Generalities

In this section, we study some attempts to recover the continuous limit. As the lattice spacing shrinks to zero, we expect that the distortions artificially introduced by the lattice disappear in some suitable limit. In particular, Euclidean invariance, which is broken by discretization, should be retrieved. Section 2 has been devoted to a careful construction of a Lagrangian in order to ensure the "best possible" continuous limit, but this is presumably not sufficient since renormalization effects may be encountered.

Let us point out that near a second-order transition point, one expects that such a limit may be reached. Under these circumstances, the correlation length becomes indeed infinite. A continuous limit may follow when re-expressing physical quantities in a relevant scale (the one pertaining to fluctuations). The local structure of the lattice is expected to be smeared. This is known to be the case for the Ising model, where the Euclidean invariance is restored at the critical point.

The renormalization group technique [48,9] allows a quantitative study of such a limit. The scaling properties of the theory lead to an equation relating a scale modification to a change in the coupling constants. In the continuous case, one obtains the Callan–Symanzik differential equations. The procedure can be extended to the discrete case using the following formulation. Let us consider an action $S[\Phi; g]$ depending on fields Φ and on coupling constants g defined on a lattice (with spacing a). Now consider another lattice, with spacing λa embedded in the preceding one, on which we define new fields Φ'. The quantities of interest are assumed to be described alternatively in terms of the new fields Φ'; i.e. we suppose that there exists a conversion functional $\Delta(\Phi, \Phi')$ relating old and new fields. A new action can be defined through

$$e^{S'[\phi']} = \int \mathscr{D}\Phi \Delta(\Phi, \Phi') e^{S[\Phi; g]}.$$ (5.1.1)

For instance, if $\{\Phi'\}$ is a subset of the original $\{\Phi\}$ obtained by selecting one site out of two ($\lambda = 2$), then $\Delta(\Phi, \Phi') = \Pi_i \delta(\Phi_i - \Phi'_i)$ where the index i runs over all selected sites. Equation (5.1.1) specifies nothing but summation over fields located on rejected sites.

The new action $S'[\Phi']$ can in general be very complex. Nevertheless, if the choice of the primitive action $S[\Phi; g]$ is sufficiently general and if Δ has been carefully chosen, it is possible that the new action S' takes the same form as the primitive one. One therefore has

$$S'[\phi'] = S[\Phi'; g'],$$

with different coupling constants g'. It is then possible to rewrite eq. (5.1.1) as

$$g' = R_\lambda(g).$$ (5.1.2)

This is the recurrence equation searched for. To familiarize oneself with this procedure, let us first apply this technique to the elementary case of a two-dimensional gauge problem. We recall that an exact solution is known (see section 4.2). Our aim is to illustrate the possibilities of the method and, in particular, to recover the existence of a "transition point" for $\beta = 0$.

5.2. Two-dimensional recurrence equation

Consider on a two-dimensional square lattice a pure gauge field dynamics. The action for an

elementary plaquette reads

$$\exp[\beta\chi(RRRR)] = \sum_r \beta_r \chi^r(RRRR),$$
(5.2.1)

where we have summed over all irreducible representations of the group, including the trivial one. Now, let us double the lattice spacing, according to fig. 14. The four elementary squares A, B, C, D are joined in one big square. New fields are defined on the new links by the product of the gauge fields on the old links as for instance $R'_{12} = R_{15}R_{52}$. The summation over the intermediate links can be performed using the formalism of sub-section 3.1.
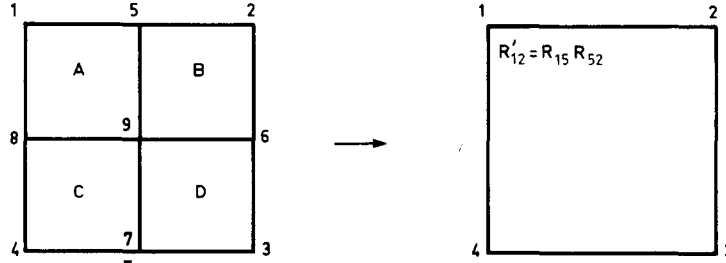
Fig. 14. Doubling the lattice spacing.

Namely,

$$\int \mathcal{D}R_{59}\mathcal{D}R_{69}\mathcal{D}R_{79}\mathcal{D}R_{89} \sum_{r_A} \sum_{r_B} \sum_{r_C} \sum_{r_D} \beta_{r_A}\beta_{r_B}\beta_{r_C}\beta_{r_D}\chi^{r_A}(R_{15}R_{59}R_{98}R_{81})$$

$$\times \chi^{r_B}(R_{52}R_{26}R_{69}R_{95})\chi^{r_C}(R_{63}R_{37}R_{79}R_{96})\chi^{r_D}(R_{74}R_{48}R_{89}R_{97})$$

$$= \sum_r \beta_r^4 \chi^r(R_{15}R_{52}R_{26}R_{63}R_{37}R_{74}R_{48}R_{81}) = \sum_r \beta_r^4 \chi^r(R'_{12}R'_{23}R'_{34}R'_{41}).$$
(5.2.2)

This allows the summation of eq. (5.1.1). We see that the result has exactly the same form as before (eq. (5.2.1)), but now the values of the new coupling constants read

$$\beta'_r = \beta_r^4.$$
(5.2.3)

More generally, if the lattice spacing a is changed to λa in a two-dimensional lattice, the interaction keeps the same form with the following change on coupling constants

$$\beta_r \rightarrow \beta'_r = \beta_r^{\lambda^2}.$$
(5.2.4)

What are the consequences of such a recurrence equation? Although the renormalization changes the coupling constants, it is clear that the qualitative properties of the system cannot change. In other words, the system remains in the same phase (short range correlations will remain short range correlations, even if the scale is doubled). Then phases are separated by the fixed points of the recurrence equation, which appears as the equation for the critical points;
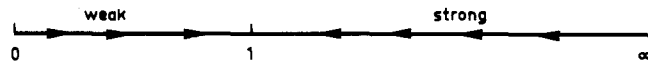
$$g_c = R_\lambda(g_c).$$
(5.2.5)

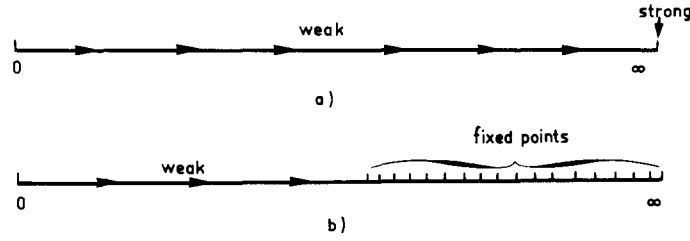Fig. 15. Phase diagram for the recurrence equation (5.2.4).

Fig. 16. Other possible phase diagrams; a) Ising model, b) Baxter model.

Figure 15 shows the characteristic shape for the type of phase diagram associated with eq. (5.2.4). There is a transition for $\beta_r^c = 1$ or 0. (Note that this is not contradictory with our analysis of subsection 4.2. The only point where *all* β_r are fixed points corresponds to $\beta = 0$, i.e. $\beta_0 = 1$ and $\beta_r = 0$ for all nontrivial representations.) The arrows show the directions of flow of coupling as the lattice spacing goes to zero.

Other types of phase diagram are possible. Figure 16 shows some possibilities: the one-dimensional Ising model, characterized by a recurrence relation $\tanh \beta' = (\tanh \beta)^\lambda$, has only disordered phase. The two-dimensional Baxter model [49] presents a continuous set of fixed points.

The conclusion of this preliminary example is that the renormalization group techniques give a first hint on the continuous limit. This limit is closely related with the critical behaviour of the system, which can be studied with a recurrence equation. We discuss now the general case of a gauge theory, where one hopes to generate the phase diagram displayed in fig. 17 for physical strong interactions.
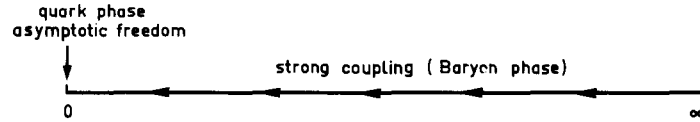
Fig. 17. The phase diagram expected for strong interactions.

5.3. Higher dimensional approximate recurrence equations

We try now to extend the previous calculation to higher dimensional lattices. The integrations over inner links was performed in eq. (5.2.2) thanks to the identity (3.1.2). The trouble is that this integration is simple if only two plaquettes interfere. In dimension greater than two, an inner link pertains to $2(d - 1)$ plaquettes and the method fails. This difficulty is circumvented when one uses an approximation proposed by Migdal [10]. Such an approximation, as we shall soon discover, may be helpful close to the lower critical dimension. The idea is to displace some interactions [50] in order to make possible the integrations. More precisely, let us consider the λ^d elementary cells $0 \le x_i < \lambda a$, which will become a single cell in the decimation procedure. The interaction involving the plaquette $(x, x + a\hat{\mu}, x + a\hat{\mu} + a\hat{\nu}, x + a\hat{\nu})$ is moved on the parallel plaquette $(y, y + a\hat{\mu}, y + a\hat{\mu} + a\hat{\nu}, y + a\hat{\nu})$, where $y_\mu = x_\mu$, $y_\nu = x_\nu$ and $y_i = 0$ otherwise (see fig. 18 for a three-dimensional example). Now the action involves only plaquettes lying on the faces of the embedded lattice with spacing λa. As each one occurs λ^{d-2} times, the primitive exponentiated action is replaced by the product over remaining plaquettes of the term

$$\exp[\lambda^{d-2}\beta\chi(RRRR)] = \left[\sum_r \beta_r \chi^r(RRRR)\right]^{\lambda^{d-2}} = \sum_r (\lambda^{d-2}\beta)_r \chi^r(RRRR). \tag{5.3.1}$$
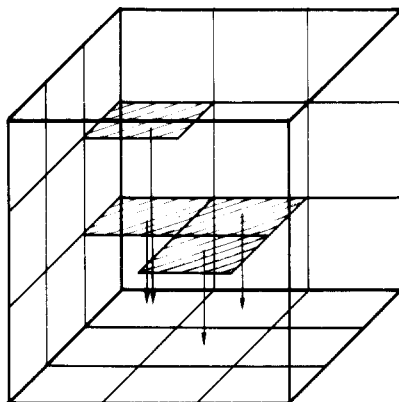
Fig. 18. Migdal's approximation.

The integration over inner links can now be achieved exactly as in the preceding subsection. The final result reads:

$$\beta'_r = (\lambda^{d-2}\beta)^{\lambda^2}_r.$$

(5.3.2)

Let us compare it with the one obtained for a usual spin system. The same derivation can be performed with the action (2.3.2) (written as $\Sigma_{(i,j)}\alpha\Phi_i\Phi_j$), invariant under the same global symmetry group. This leads to the approximate equation

$$\alpha'_r = (\lambda^{d-1}\alpha)^{\lambda}_r.$$

(5.3.3)

The comparison with (5.3.2) shows that a $d/2$-dimensional spin system, scaled by a factor λ^2, obeys exactly the same reccurrence equation. For instance, the Z_2 gauge system is closely related to the Ising model, and the four-dimensional $U(N)$ gauge systems to the corresponding two-dimensional classical spin models (or equivalently non linear σ-models). If Migdal's analysis is correct, one might therefore be able to transpose all known results on these global systems for the gauge systems with double dimension.

Let us analyse in more detail the recurrence equation, first in the Z_2 case. There are only two irreducible representations for this group; the couplings are $\beta_0 = \cosh\beta$, $\beta_1 = \sinh\beta$, and Migdal's equation reads

$$\beta'_0 = (\cosh\lambda^{d-2}\beta), \qquad \beta'_1 = (\sinh\lambda^{d-2}\beta)^{\lambda^2}.$$

Of course, only the ratio of the two couplings has a physical meaning (other dependences can be reabsorbed in a field multiplicative renormalization constant) and one finally obtains

$$\tanh\beta' = (\tanh\lambda^{d-2}\beta)^{\lambda^2}.$$

(5.3.4)

As d goes to 2, the fixed point lies near β infinite. More precisely, in this limit, the fixed point behaves as $\beta_c = 1/(d-2)$. For $d < 2$, the critical point disappears entirely. The lower critical dimension d_c is in this case equal to 2. This analysis is perfectly parallel to the similar one for the Ising model the critical dimension of which is equal to 1. Let us emphasize that all the results we have obtained for the Z_2 gauge system is the counterpart of the corresponding result for the Ising model in half-dimension.

The same analysis can be carried for a continuous Lie group. The global symmetric spin system presents in this case [10, 51, 52] a lower critical dimension of two. Therefore the gauge system shows

a critical dimension of 4. One is thus in the case of fig. 17, as one hopes for. But what about the U(1) case, corresponding to electrodynamics? In this special case, all proofs of the existence of such a behaviour fail. The XY model (i.e. the global symmetric spin version) has probably [52] a non trivial transition at $d = 2$, even though the order parameter cannot have a non vanishing value in the "ordered" phase. If Migdal's correspondence is reliable, this is physically sound since we observe free electrons and since quantum electrodynamics is not asymptotically free. Therefore these arguments nicely fit together.

Of course the above analysis is only very preliminary and does not give us a reliable method to reach the continuous limit in a non perturbative way. It is clear that the infinite renormalizations of the latter should appear somewhere even though asymptotic freedom of non abelian gauge fields should perhaps simplify the treatment. Also some gauge fixing mechanism, avoided by lattice quantization has somehow to be reintroduced if color non invariant concepts are to be used at short distances. This shows that, although not hopeless, the task of a predictive solution to the confinement problem remains unfortunately a task for the future.

Appendix A. Some definitions and results in combinatorial topology

We collect here some definitions and results of combinatorial topology. We omit the proofs and refer to the literature [54] for further details.

A finite set of mutually disjoint convex polyhedral domains is called a *polyhedral complex* if every face of every domain is also an element of the complex. The maximum dimension of the domains is the *dimension* of the complex. A complex is said *pure* if each constituent domain of non maximum dimension is a face of another constituent domain. We used in this paper one- and two-dimensional complexes built with the nodes, links and plaquettes of a hypercubical lattice.

In a pure one-dimensional complex, there are *regular* points (a neighbourhood of which is homeomorphic to a line) and *singular* points (divided into *end points* and *branch points*). It can be partitioned in *regular components*, which are either a simple closed curve (homeomorphic to a circle) with at most one singular point, or a simple arc (homeomorphic to a segment) with exactly two singular points, its extremities. This type of topology occurs in usual field theory (Feynman diagrams).

Regular points of a pure two-dimensional complex have a neighbourhood homeomorphic to a plane. Non isolated singular points form a pure one-dimensional complex called the *contour*. The *boundary* is the set of contour points, a neighbourhood of which is homeomorphic to a half-plane. If the neighbourhood of a contour point is homeomorphic to n half-planes bounded by the same line, this point belongs to a *n-fold branch line*. Finally, the complex is partitioned in *regular components*; they are subcomplexes such that all their singular points are singular points of the complex and which cannot be disconnected by a cut along singular lines. All singular points of a *surface* are boundary points, and regular components are always surfaces. A *closed surface* has only regular points. A surface is indeed a closed surface with holes.

Due to this decomposition in regular components, the classification of two-dimensional complexes derives from the classification of closed surfaces. These may be:

– orientable: a sphere with $2p$ holes closed with p first kind handles (fig. 19) ($p = 0$ (sphere), 1 (torus), 2, ...).
– non orientable: a sphere with p holes closed with Möbius band (fig. 20) ($p = 1$ (projective plane), 2 (Klein bottle), 3, ...).
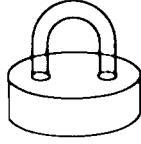
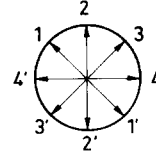Fig. 19. Closing two holes with a first kind handle.

Fig. 20. The point identification allowing to close a hole with a Möbius band.

Appendix B. Embedding two-dimensional diagrams on lattices

A graph is a set of plaquettes of a given lattice, while a diagram is the topological description of the complex formed by a graph. This appendix is devoted to the calculation of the number of configurations of a given diagram, i.e. the number of graphs associated to this diagram.

We restrict ourselves to a finite lattice of N nodes with periodic boundary conditions. Due to translation invariance, the number of configurations, denoted by $\{D\}$ for a diagram D, is a polynomial in N, which vanishes for $N = 0$, and which has a degree equal to the number of connected parts in the diagram. We are also interested by the coefficient $[D]$ of order 1 in this polynomial, the reduced number of configurations.

The number of configurations for disconnected diagrams is given by a recurrence formula. By successive embeddings of two diagrams D_1 and D_2 on the same lattice, one gets the identity

$$\{D_1\} \cdot \{D_2\} = \sum_{D=D_1 \cup D_2} n_D \{D\}$$ (B.1)

where the summation runs over all different diagrams D which can be decomposed as $D_1 \cup D_2$ in n_D different ways. Only one of these diagrams is composed of all the disconnected pieces of D_1 and D_2, and this allows a recursive calculation of the number of configurations for all disconnected diagrams. For instance,

$$\{\square\}^2 = 2\{\square\ \square\} + 2\{\square\square\} + \{\square\}$$

$$\{\square\}\{\square\square\} = \{\square\ \square\square\} + 2\{\square\square\} + 2\{\boxed{\square}\} + 2\{\square\square\square\} + 3\{\square\!\!\diamondslash\}.$$

By taking the term linear in N, formula (B.1) is easily extended to the computation of reduced number of configurations:

$$\sum_{D=D_1 \cup D_2} n_D [D] = 0.$$ (B.2)

Such a method can be refined to express these quantities in terms of these pertaining to connected diagram which can be separated by a cut along one link only. Let us distinguish a link in the diagram and suppose that there are k links in the diagram playing the same topological role. The number of different embeddings which map the distinguished link at a given place on the lattice is

$$(Nd)^{-1} k\{D\} = k[D]/d.$$ (B.3)

If D can be decomposed into $D_1 \cup D_2$, both sharing the distinguished link, in m_D ways, one has

$$\frac{k_1[D_1]}{d} \frac{k_2[D_2]}{d} = \sum_{D=D_1 \cup D_2} m_D \frac{k[D]}{d}.$$ (B.4)

For instance,

$$\frac{4[\square]^2}{d} = 2\frac{[\boxed{\square}]}{d} + \frac{4[\boxed{\square}]}{d}, \qquad \frac{[\boxed{\square}]}{d}\frac{[\boxed{\square}]}{d} = 2\frac{[\boxed{\square}]}{d} + 3\frac{\left[\boxed{\square\diamondsuit}\right]}{d}$$
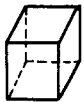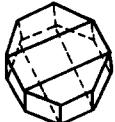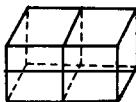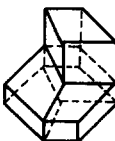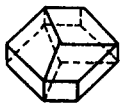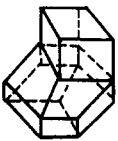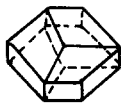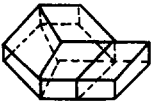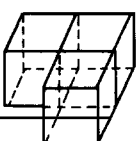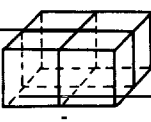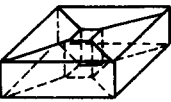
and so on.

We do not know of any general rule to write the number of configurations for connected, irreducible diagrams. We give in table 2 a list for lower order diagrams, which have to be computed case by case.

Table 2a.
Irreducible diagrams up to 5th order.

$$\left[\;\right] = \frac{1}{2}\, d(d-1)$$

$$\left[\;\right] = \frac{4}{3}\, d(d-1)(d-2)$$

$$\left[\;\right] = \frac{1}{2}\, d(d-1)(4d^2-16d+17)$$

$$\left[\;\right] = 2d(d-1)(d-2)$$

$$\left[\;\right] = \frac{1}{2}\, d(d-1)(d-2)$$

$$\left[\;\right] = d(d-1)(d-2)$$

$$\left[\;\right] = \frac{4}{5}\, d(d-1)(d-2)(4d^2-18d+23)$$

$$\left[\;\right] = 4d(d-1)(d-2)(2d-5)$$

$$\left[\;\right] = 2d(d-1)(d-2)(2d-5)$$

$$\left[\;\right] = 4d(d-1)(d-2)(2d-5)$$

$$\left[\;\right] = 4d(d-1)(d-2)(2d-5)$$

Table 2b.
Simple closed surfaces, up to 16th order.

$$\left[\ \right] = \frac{1}{6} d(d-1)(d-2)$$

$$\left[\ \right] = d(d-1)(d-2)(d-3)$$

$$\left[\ \right] = \frac{1}{2} d(d-1)(d-2)(2d-5)$$

$$\left[\ \right] = 16 d(d-1)(d-2)(d-3)^2$$

$$\left[\ \right] = \frac{4}{3} d(d-1)(d-2)(d-3)$$

$$\left[\ \right] = 8 d(d-1)(d-2)(d-3)^2$$

$$\left[\ \right] = \frac{1}{3} d(d-1)(d-2)(d-3)$$

$$\left[\ \right] = 16 d(d-1)(d-2)(d-3)^2$$

$$\left[\ \right] = 2 d(d-1)(d-2)(4d^2-20d+25)$$

$$\left[\ \right] = \frac{1}{2} d(d-1)(d-2)(4d^2-24d+37)$$

$$\left[\ \right] = \frac{1}{2} d(d-1)(d-2)(4d^2-20d+25)$$

$$\left[\ \right] = 8 d(d-1)(d-2)(d-3)$$
(torus)

Appendix C. Gauge groups and lattice theory

C.1. Discrete, Abelian group Z_2

The group $Z_2 = \{1, -1\}$ has two *real, one-dimensional* representations.

In the strong coupling expansion, we therefore deal with unoriented plaquettes, and the sources reduce to real numbers K. The invariant integration is the mean over the field values ± 1, and the unperturbed free energy per link reads

$$u(K) = \text{Ln}\left(\frac{1}{2}\sum_{R=\pm 1} \exp(KR)\right) = \text{Ln} \cosh K. \tag{C.1}$$

The system has been extensively studied in subsection 4.3. Let us nevertheless illustrate the expansion of Z (subsection 3.3). We have computed F up to 16th order. In order to perform this calculation, we have to consider all the two-dimensional polyhedral complexes. These are classified in appendix A and listed in appendix B with their number of configurations. Furthermore, each constituent square must be dressed with one, two, three,... plaquettes. A link belonging to p plaquettes yields a term

$$\frac{d^p}{dK^p} \exp u(K) = \begin{cases} 0 & \text{if } p \text{ is odd} \\ 1 & \text{if } p \text{ is even.} \end{cases} \tag{C.2}$$

For instance, the single square, which can be mapped $\frac{1}{2}Nd(d-1)$ times on the lattice, can be dressed with 2, 4, 6, plaquettes and yields a term

$$N\frac{d(d-1)}{2} \times \left(\frac{\beta^2}{2!} + \frac{\beta^4}{4!} + \frac{\beta^6}{6!} + \cdots\right).$$

The lowest order cube diagram contributes a factor $\frac{1}{6}Nd(d-1)(d-2)\beta^6$. The next non vanishing contribution is obtained by dressing one of the six faces by three plaquettes and yields $\frac{1}{6}Nd(d-1)(d-2) \times 6\,\beta^8/3!$. Finally, one gets, after taking the logarithm and assuming the usual infinite volume (that is infinite N) limit (i.e. the coefficient of N in Z):

$$\frac{F}{Nd(d-1)} = \frac{1}{4}\beta^2 - \frac{1}{24}\beta^4 + \left(\frac{1}{6}d - \frac{29}{90}\right)\beta^6 + \left(-\frac{1}{3}d + \frac{3343}{5040}\right)\beta^8 + \left(d^2 - \frac{184}{45}d + \frac{118471}{28350}\right)\beta^{10}$$

$$+ \left(-\frac{8}{3}d^2 + \frac{121153}{11340}d - \frac{20022781}{1871100}\right)\beta^{12} + \left(10d^3 - \frac{208}{3}d^2 + \frac{935561}{5670}d - \frac{5647451354}{42567525}\right)\beta^{14}$$

$$+ \left(-\frac{74}{3}d^3 + \frac{120761}{840}d^2 - \frac{345869921}{1247400}d + \frac{3612986481191}{20432412000}\right)\beta^{16} + O(\beta^{18}). \tag{C.3}$$

In this case, the original expansion of subsection 3.1 appears as a series in $\beta_1 = \tanh \beta$ and only involves closed surfaces. It is therefore easier to compute, but of course gives the same result.

C.2. Discrete, Abelian group Z_n, $n > 2$

The group $\{\exp(2ik\pi/n); k = 0, \ldots n-1\}$ has n complex, one-dimensional representations, namely $\{\exp(2ikr\pi/n)\}$ for $r = 0$ (trivial), 1 (fundamental), $2, \ldots, n-1$. Representations r and $n-r$ are adjoint. Orientation must now be taken into account. The source K is a complex number and we have

$$u(K) = \mathrm{Ln} \left(\frac{1}{n} \sum_{k=0}^{n-1} \exp[K \, \exp(2ik\pi/n) + K^* \exp(-2ik\pi/n)] \right) = \mathrm{Ln} \sum_{\substack{\lambda, \, \mu \geq 0 \\ (\lambda - \mu)/n \text{ integer}}} \frac{K^\lambda K^{*\mu}}{\lambda! \mu!}. \qquad (C.4)$$

As a consequence, an oriented link bordering λ plaquettes with the same orientation and μ with the reverse one yields

$$\frac{\partial^{\lambda+\mu}}{\partial K^\lambda \partial K^{*\mu}} \exp u(K) = \begin{cases} 1 & \text{if } (\lambda - \mu)/n \text{ is integer} \\ 0 & \text{otherwise.} \end{cases} \qquad (C.5)$$

Hence selection rules arise for dressing the diagrams. Note that non-orientable surfaces are permitted only if n is even $(r = \frac{1}{2}n$ is indeed a self-adjoint representation). Apart from these complications, the calculations are the same as before.

C.3. Unitary U(n) and special unitary SU(n) groups

The unitary group U(n) can be parametrized as follows [55]: $R = UDU^{-1}$, with $D = $ diag. matr. $\{\exp(2i\pi\varphi_1), \ldots, \exp(2i\pi\varphi_n)\}$ belonging to an n parameter abelian subgroup Λ, and U to SU(n)/Λ. With this parametrization, the invariant measure over Λ reads

$$\prod_{1 \leq i < j \leq n} [1 - \cos(\varphi_i - \varphi_j)] \cdot \prod_{i=1}^{n} \frac{d\varphi_i}{2\pi}.$$

The fundamental character depends only on the angles φ_i and reduces to $\mathrm{Tr}\, D = \sum_{i=1}^{n} \exp(2i\pi\varphi_i)$. Finally the reduction to SU(n) is obtained by imposing the constraint $\sum_{i=1}^{n} \varphi_i = 0 \pmod{2\pi}$, which is easily realized by a δ function in the measure.

A first example is U(1) (or SO(2)). This group is abelian; the fundamental representation is complex and one dimensional. Hence

$$u(K) = \mathrm{Ln} \int_0^{2\pi} \exp(K \, e^{2i\pi\varphi} + K^* \, e^{-2i\pi\varphi}) \frac{d\varphi}{2\pi} = \mathrm{Ln}\, I_0(2|K|), \qquad K \text{ complex number} \qquad (C.6)$$

where I_n is the modified Bessel function of order n.

Graphs are oriented, $u(K)$ depends only on KK^* and thus there is a selection rule forbidding links not connected to equal number of plaquettes with reverse orientations.

As a second particular case, let us discuss SU(2) (or SO(3)). Although the representations are complex, the characters are real. Then only unoriented plaquettes are to be considered. The source K is a 2×2 real matrix, but $u(K)$ depends only on its determinant

$$u(K) = \ln \frac{I_1(2\eta)}{\eta}, \text{ with } \eta^2 = \det K. \qquad (C.7)$$

With the other groups, we are in the general case, with oriented graphs and an $n \times n$ complex matrix source K.

References

[1] A review on gauge theories can be found in E.S. Abers and B.W. Lee, Phys. Reports 9C (1973) 1.
[2] See e.g. H.D. Politzer, Phys. Reports 14C (1974) 130.
[3] O.W. Greenberg, Phys. Rev. Lett. 13 (1964) 598;
    A modern discussion can be found in H. Fritzsh, M. Gell-Man and H. Leutwyler, Phys. Lett. 47B (1973) 365.

[4] K.G. Wilson, Phys. Rev. D10 (1974) 2445.
[5] A.M. Polyakov, unpublished.
[6] R. Balian, J.M. Drouffe and C. Itzykson, Phys. Rev. D10 (1974) 3376.
[7] R. Balian, J.M. Drouffe and C. Itzykson, Phys. Rev. D11 (1975) 2104.
[8] J. Kogut, D.K. Sinclair and L. Susskind, Nucl. Phys. B114 (1976) 199.
[9] L.P. Kadanoff, Rev. Mod. Phys. 49 (1977) 267.
[10] A.A. Migdal, Zh. E.T.F. 69 (1975) 810; 69 (1975) 1477.
[11] Berezin, The method of second quantization (Academic press, N.Y., 1966).
[12] J. Kogut and L. Susskind, Phys. Rev. D11 (1975) 395.
[13] W.A. Bardeen and R.B. Pearson, Phys. Rev. D14 (1976) 547.
[14] S.D. Drell, M. Weinstein and S. Yankielowicz, Phys. Rev. D14 (1976) 487; D14 (1976) 1627.
[15] K.G. Wilson, Relativistically invariant lattice theories, Coral Gable conference (January 76), Cornell preprint CLNS 327.
[16] K.G. Wilson, Erice lecture notes (1975).
[17] A. Casher, Y. Aharonov and L. Susskind, Phys. Rev. D8 (1973) 440.
[18] C.N. Yang and R. Mills, Phys. Rev. 96 (1954) 191, see also ref. [1].
[19] M. Creutz, Phys. Rev. D15 (1977) 1128.
[20] A review of statistical models can be found in H.E. Stanley, Introduction to phase transitions and critical phenomena (Clarendon press, Oxford, 1971).
[21] R. Balian, J.M. Drouffe and C. Itzykson, Phys. Rev. D11 (1976) 2098.
[22] P.W. Higgs, Phys. Rev. Lett. 12 (1964) 132.
[23] J. Schwinger, Phys. Rev. 125 (1962) 397; 128 (1962) 2425.
[24] J. Lowenstein and A. Swieca, Ann. Phys. 68 (1971) 172.
[25] A. Casher, J. Kogut and L. Susskind, Phys. Rev. Lett. 31 (1973) 792; Phys. Rev. D10 (1974) 732.
[26] S. Coleman, Harvard preprint (1976);
     R. Jackiw, L. Susskind and S. Coleman, Ann. Phys. 93 (1975) 267.
[27] M. W. Roth, Phys. Rev. D15 (1977) 1084.
[28] G.'t Hooft, Nucl. Phys. B75 (1974) 461.
[29] C.G. Callan, N. Coote and D.J. Gross, Phys. Rev. D13 (1976) 1649.
[30] M. Einhorn, Phys. Rev. D14 (1976) 3451.
[31] C.P. Korthals Altes, Marseilles Colloquium on Lagrangian field theory (June 1974).
[32] C. Rebbi, Phys. Reports 12C (1974) 1.
[33] Y. Nambu, Lectures for the Copenhagen Summer Symp. 1970, unpublished.
[34] H. Harari, Phys. Rev. Lett. 22 (1969) 562;
     J.L. Rosner, Phys. Rev. Lett. 22 (1969) 689.
[35] S. Okubo, Phys. Lett. 5 (1963) 165;
     G. Zweig, CERN, unpublished;
     J. Iizuka, Progr. Theor. Phys. Suppl. 37–38 (1966) 21.
[36] H.G. Dosch and V.F. Müller, Phys. Rev. D12 (1975) 3343.
[37] C. Domb, Adv. Phys. 9 (1960) 149; 19 (1970) 339.
[38] F. Englert, Phys. Rev. 129 (1963) 567.
[39] C. Bloch, Studies in quantum statistical mechanics, Vol. 3, eds. J. De Boer and G.E. Uhlenbeck (North-Holland, Amsterdam, 1965) p. 1.
[40] M.E. Fisher and D.S. Gaunt, Phys. Rev. 133A (1964) 224;
     R. Abe, Progr. Theor. Phys. 47 (1972) 62.
[41] C. Bloch, Nucl. Phys. 6 (1958) 329.
[42] T. Banks, L. Susskind and J. Kogut, Phys. Rev. D13 (1976) 1043.
[43] A. Caroll, J. Kogut, D.K. Sinclair and L. Susskind, Phys. Rev. D13 (1976) 2270.
[44] J. Shimegitsu and S. Elitzur, Phys. Rev. D14 (1976) 1988.
[45] R.E. Peierls, Phys. Rev. 54 (1938) 918.
[46] F.J. Wegner, J.M.P. 12 (1971) 2259.
[47] R.B. Griffiths, J.M.P. 8 (1967) 478;
     D.G. Kelly and S. Sherman, J.M.P. 9 (1968) 466.
[48] K.G. Wilson and J. Kogut, Phys. Reports 12 (1974) 75.
[49] R. Baxter, Ann. Phys. (N.Y.) 70 (1972) 193.
[50] L. Kadanoff, A. Houghton and M.C. Yalabik, J. Stat. Phys. 14 (1976) 171;
     L. Kadanoff, Phys. Rev. Lett. 34 (1975) 1005.
[51] E. Brezin and J. Zinn-Justin, Phys. Rev. Lett. 36 (1973) 691.
[52] A.M. Polyakov, Phys. Lett. 59B (1975) 79.
[53] J. Zittartz, Z. Phys. 206 (1971) 465.
[54] P.S. Aleksandrov, Combinatorial Topology, Vol. 1 (Graylock press, 1956).
[55] H. Weyl, The classical groups (Princeton, 1946).