THERMODYNAMICS AND TWO-DIMENSIONAL LATTICE GAUGE MODELS*

By J. Jurkiewicz

Institute of Physics, Jagellonian University, Cracow

and K. Zalewski

Institute of Nuclear Physics, Cracow**

(Received January 12, 1983)

The gauge theory with the gauge group $U(N \to \infty)$ is solved on a two-dimensional lattice. The single plaquette action used depends on L parameters, where L is an arbitrary integer, and thus results for a wide class of variant actions may be compared. A rich structure of second order and third order phase transitions appears. Besides the exact analytic solution a thermodynamical discussion clarifying the qualitative features of the results is given.

PACS numbers: 11.15.-q, 11.15.Ha

1. Introduction

We present the solution of a problem in lattice gauge theory. The solution is obtained using established mathematical methods and could be presented without reference to thermodynamics. When looking at the formulae, however, one easily notices that many of the necessary concepts and quantities have thermodynamic interpretations. Using simple theorems from thermodynamics it is possible not only to interpret the formulae, but often also to foresee results quite difficult to obtain by formal calculations. Thus the analysis presented here, besides its intrinsic interest for lattice gauge theory, provides an instructive example of the blend between statistical thermodynamics and quantum field theory so characteristic for modern developments in both fields. We begin with some general remarks about lattice formulations of field theories.

2. Lattice formulation of quantum field theory as a way to regularization

The lattice formulation of quantum field theory can be considered as a method of introducing ultraviolet regularization. In order to illustrate this point let us consider the Casimir effect in one space and one time dimension. The arrangement is shown in Fig. 1.

---

* Presented by one of the authors (K. Z.) at the Smolenice Conference, November 1982.

** Address: Instytut Fizyki Jądrowej, Kawiory 26a, 30-055 Kraków, Poland.

Two "metal plates" are at a distance L from each other. Any oscillation, which could be excited in the region $0 \leqslant x \leqslant L$, must have zeros of the amplitude at $x = 0$ and at $x = L$. Therefore, the longest wave possible is the one shown in the upper part of Fig. 1, while the second longest is that in the lower part of the figure. All intermediate wave lengths

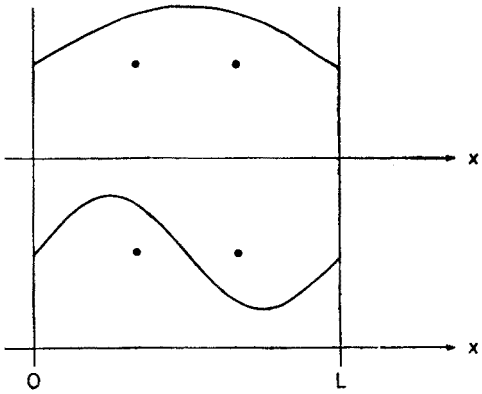

Fig. 1. Examples of allowed oscillations contributing to the Casimir effect. The black dots in each part of the figure form a lattice ($N = 2$)

are forbidden. Since according to quantum field theory a zero point energy $\frac{1}{2}\omega$ is associated with any possible oscillation of frequency ω, the lowest energy state of the region between the two plates has energy

$$E_0(L) = \tfrac{1}{2} \sum_{n=1}^{\infty} \omega_n, \tag{2.1}$$

where the index n labels the possible oscillations, and the dependence of this energy on L is explicitly noted. For the three-dimensional case (and for electromagnetic oscillations) this sum has been evaluated by Casimir and the result suitably regularized and renormalized agrees very well with experiment. For the one-dimensional case

$$\omega_n = \frac{2\pi}{\lambda_n} = \frac{\pi n}{L}, \tag{2.2}$$

therefore

$$E_0(L) = \frac{\pi}{2L} \sum_{n=1}^{\infty} n. \tag{2.3}$$

This cannot be directly used, but regularization and renormalization are required. Note that the divergence in expression (2.3) is of the ultraviolet type, i.e. results from the possibility of arbitrarily short waves. There are many methods of regularizing the sum in formula (2.3). The shortest way is to use the definition of Riemann's dzeta function

$$\zeta(s) = \sum_{n=1}^{\infty} n^s \tag{2.4}$$

valid for $s < -1$, and to identify the sum in expression (2.3) with $\zeta(-1) = -1/12$. Here, however, we are interested with lattice regularization. Let us introduce lattice points at

$$x = \frac{k}{N+1} L, \quad k = 1, \ldots, N \tag{2.5}$$

and convert the differential equation for the amplitude of the oscillations

$$\Box A = 0 \tag{2.6}$$

into a difference equation in the space variable:

$$\frac{d^2 A(k)}{dt} - a^{-2}[A(k+1) - 2A(k) + A(k-1)] = 0, \quad k = 1, \ldots, N, \tag{2.7}$$

with the boundary conditions

$$A(0) = A(N+1) = 0. \tag{2.8}$$

This approach, where time remains continuous and a lattice is introduced for the space dimensions only, is known as the Hamiltonian method. The solutions periodic in time read for equation system (2.7)

$$A_n(k) = e^{i\omega_n t} \sin \nu_n k, \quad n = 1, \ldots, N, \tag{2.9}$$

where

$$\nu_n = \frac{\pi a n}{L} \tag{2.10}$$

and the eigenfrequencies are

$$\omega_n = \frac{2}{a} \sin\left(\frac{\nu_n}{2}\right). \tag{2.11}$$

It is important that there are only N independent solutions. The solution for $n = N+1$ vanishes identically, while that for $n = N+2$ reproduces that for $n = 1$ and so on. Therefore, the divergent sum (2.1) is replaced by the finite sum

$$E_0(L; a) = \frac{1}{a} \sum_{n=1}^{N} \sin \frac{\pi n}{2(N+1)} = \frac{1}{2a}\left[\operatorname{ctg} \frac{\pi a}{4L} - 1\right]. \tag{2.12}$$

In order to go to the continuum limit, the right hand side of this relation should be expanded in powers of a:

$$E_0(L; a) = \frac{2L}{\pi a^2} - \frac{1}{2a} - \frac{\pi}{24L} + O\left(\frac{a^2}{L^3}\right). \tag{2.13}$$

The terms not written out explicitly vanish in the continuum limit $a \to 0$. Formula (2.13) has the required regularized form. The third term is the correct finite answer (cf. e.g. Ref.

[1]). The first two infinite terms must be removed by renormalization. In order to eliminate the first term let us notice that the quantity of interest is not just $E_0(L)$, but the difference between $E_0(L)$ and the energy contained in the relevant region $0 \leqslant x \leqslant L$, when there are no plates. Denoting this difference by $\Delta E_0(L)$, we have in the limit $a \to 0$, $\mathcal{N} \to \infty$:

$$E_0(L) = E_0(L) - \frac{1}{\mathcal{N}} E_0(\mathcal{N}L) = -\frac{\pi}{24L} + \text{const.} \tag{2.14}$$

The constant is formally infinite, but it is not measurable and can be put equal zero by convention. Thus the required finite result is obtained.

3. Transition to the continuum limit, when the Monte-Carlo method is being used

In the preceding section the transition to the continuum limit was simple, because an analytic expression for the quantity of interest was known. Usually, however, the lattice calculations are performed using the Monte-Carlo method and then only numerical results are available. All dimensional quantities are obtained in units of powers of the lattice spacing a, and it is by no means obvious, how the transition to the continuum limit should be made. In such cases the renormalization group provides us with the necessary tools. For an SU(N) gauge group the following formula relates the coupling constant to the lattice spacing

$$\Lambda a = \left(\frac{11x}{3}\right)^{-\frac{51}{121}} \exp\left(-\frac{3}{22x}\right)(1 + O(x)), \tag{3.1}$$

where

$$x = \frac{g^2 N}{4\pi^2}. \tag{3.2}$$

The gauge theory scale Λ cannot be calculated from the theory. Even without knowing it, however, it is clear from the formula that the continuum limit $a \to 0$ corresponds to the weak coupling limit, $x \to 0$, of the theory on the lattice. It is remarkable that however large the continuum coupling constant is, the continuum limit of the lattice theory is always obtained for the lattice coupling constant tending to zero.

Any dimensional quantity can be calculated in units of Λ to a suitable power. For instance, the so called string tension σ has dimension cm^{-2}. Thus the dimensionless parameter is

$$\frac{\sigma}{\Lambda^2} = \frac{\sigma a^2}{(\Lambda a)^2}. \tag{3.3}$$

The numerator (σa^2) equals σ in units of a^{-2}. Therefore it can be obtained by the Monte-Carlo method. The denominator $(\Lambda a)^2$ can be obtained from formula (3.1). Therefore, for given g^2 the ratio σ/Λ^2 is calculable on a lattice. For a sufficiently small it should not depend on a and a smooth extrapolation to $a = 0$ should give the continuum limit.

The scale Λ is not calculable, and in the continuum theory depends on the renormalization scheme chosen. This corresponds in the lattice theory to the choice of the action on the lattice (cf. following section). The dependence of Λ on the renormalization scheme is calculable. It is enough to evaluate any measurable quantity in the two schemes. Each result will depend on the corresponding Λ. Comparing the two results, which should be the same, one finds the relation between the $\Lambda - s$. Thus e.g. calculating the vacuum to vacuum transition amplitude

$$Z = \int \mathcal{D}U e^{-S(U)}, \tag{3.4}$$

one finds the relation between the scales Λ for lattice gauge theory with Wilson's action and the continuum theory with the Pauli-Villars renormalization scheme, both for the SU(N) gauge theory [2, 3]

$$\Lambda_{\mathrm{W}} = 0.023680 \exp\left(-\frac{3\pi^2}{11N^2}\right) \Lambda_{\mathrm{PV}}. \tag{3.5}$$

The application of formula (3.1) in practical calculations requires some care. For g^2 too small the Monte-Carlo method becomes very inefficient. On the other hand for g^2 too large, formula (3.1) may be invalidated by the presence of a phase transition or some other nonperturbative effects. The existence of a window with g^2 neither too small nor too large is a necessary and not obviously satisfied condition for the applicability of the approach presented here. In particular it is quite possible that this window is present for some choices of the lattice action and absent for others [4]. This is one of the reasons why it is important to study the dependence of lattice gauge theories on the choice of the action on the lattice.

4. Variant actions on a lattice

A lattice consists of plaquettes. For a gauge group U(N), to every link of every plaquette a unitary matrix U_i is ascribed. The index i labels the links. The product of the four matrices U_i corresponding to the four links forming a plaquette (cf. Fig. 2) is the building block

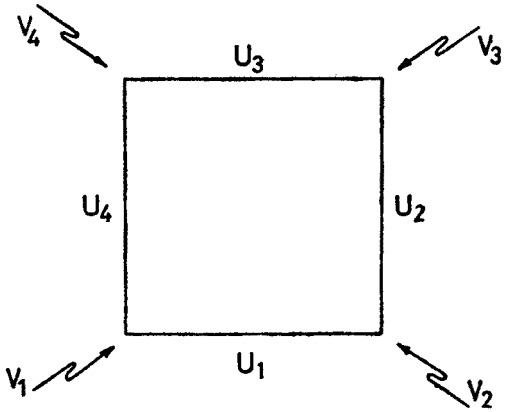

Fig. 2. Plaquette P

for the single plaquette action. A useful notation is

$$U_P = U_1 U_2 U_3 U_4. \tag{4.1}$$

The subscript P denotes here the plaquette delimited by the four links 1, 2, 3, 4. The trace of U_P is gauge invariant in the sense that for any four unitary matrices $V_1, \ldots, V_4$

$$\mathrm{Tr}\, U_P = \mathrm{Tr}\,(V_1^+ U_1 V_2 V_2^+ U_2 V_3 V_3^+ U_3 V_4 V_4^+ U_4 V_1). \tag{4.2}$$

Also the trace of an arbitrary polynomial in U_P and U_P^+ is gauge invariant.

The action for the whole lattice is usually assumed equal to the sum of the single plaquette actions corresponding to all the plaquettes forming the lattice

$$S = \sum_P S_P. \tag{4.3}$$

The single plaquette action S_P should be a gauge invariant function of the corresponding matrix U_P. Here, however, there are many possibilities. Thus Wilson's action [5] is defined by

$$g^2 S_P = \mathrm{Tr}\,(U + U^+ - 2). \tag{4.4}$$

Here and in the following the subscript P of U_P is omitted. Manton's action [6] is

$$g^2 S_P = -\mathrm{Tr}\, X^2, \tag{4.5}$$

where

$$U = e^{iX}. \tag{4.6}$$

There is an infinity of other choices. There are only two restrictions. Firstly, the action S_P should be gauge invariant — this is always satisfied, when S_P is the trace of a power series in U and U^+. Secondly, the action S_P should have the correct "naive continuum limit", i.e. relation (4.5) should hold in the limit $X \to 0$. The hypothesis that any lattice action, which satisfies these two conditions, yields in the continuum limit the same field theory is known under the name of universality.

For $a \neq 0$, or equivalently for $X \neq 0$, lattice gauge theories, which are equivalent according to the universality hypothesis, may behave quite differently. Thus for instance, the gauge theory on a two dimensional lattice for the gauge group U(N) and $N \to \infty$ has a phase transition, when Wilson's action is used [7], and has no phase transition, when Manton's action is used [8]. As explained in the preceding section, such differences may be very important for practical calculations, even if the continuum limits are identical in the two approaches. In order to get some insight into the variety of the possible phase structures, we will study the gauge theory on a two-dimensional lattice for the gauge group U(N) with $N \to \infty$ assuming

$$S_P = \sum_{l=1}^{L} \frac{\beta_l}{2l} \mathrm{Tr}\,(U^l + (U^+)^l). \tag{4.7}$$

The L real constants β_l are free. Since we expect complicated phase structures, the Monte-Carlo method cannot be used. In order to make an analytic solution possible, we had to limit our discussion to a two-dimensional lattice and to the comparatively simple gauge group U(N) with $N \to \infty$. For $L = 1$ action (4.7) reduces to Wilson's with $\beta = \beta_1/2$ and with an irrelevant constant omitted.

5. *The vacuum structure*

All the necessary information about the lattice system is contained in the partition function Z or in the function F defined by

$$Z = e^F = \int \mathscr{D}U e^{-\tilde{S}(U)}. \tag{5.1}$$

Here $\tilde{S}(U)$ is the total action of the lattice for given matrices U_i on all the links and $\mathscr{D}U$ is an (invariant) integration over all possible matrices on all the links. In two dimensions expression (5.1) can be replaced by a much simpler one. Consider the plaquette shown in Fig. 2. By a suitable choice of gauge (here of the matrices V_1 and V_2) one can always replace the matrices U_2 and U_4 by unit matrices. This corresponds to the temporal gauge in the continuum theory. Moreover, since the integration $\mathscr{D}U$ is invariant, we can replace the matrix U_3 by $U_1^{-1}U_3$. Then U_P reduces to U_3. It is easy to convince oneself that a similar transformation can be performed on all the plaquettes simultaneously, if we ignore possible problem at the boundary of the lattice. Since the total number of plaquettes $\mathscr{N}^2$ is assumed very large, we indeed can forget about the O($\mathscr{N}$) boundary plaquettes. One also can check that this transformation cannot be realized on a lattice in more than two dimensions. After the transformation, the integral (5.1) factorizes into (approximately) $\mathscr{N}^2$ identical integrals, each over one matrix. Thus we can calculate first for a single plaquette P

$$Z_P = e^{F_P} = \int \mathscr{D}U_P e^{\tilde{S}_P(U_P)} \tag{5.2}$$

and then, if necessary, use the relations

$$Z = (Z_P)^{\mathscr{N}^2}, \quad F = \mathscr{N}^2 F_P. \tag{5.3}$$

Since, however, Z_P and F_P contain all the information we need, we will use only formula (5.2) omitting in the following the subscripts P.

Using a well-known theorem of Weyl one can reduce the integral (5.2) to an N-fold integral over the eigenvalues of U_P (cf. e.g. Ref. [7]). Since U_P is a unitary matrix, each of its eigenvalues can be written as $\exp(i\varphi)$ and

$$Z = \int_0^{2\pi} d^N\varphi \exp\left(\sum_{i=1}^{N} \tilde{S}(\varphi_i) + \tfrac{1}{2}\sum_{j\neq 1}^{N} \ln \sin^2\left(\frac{\varphi_i - \varphi_j}{2}\right)\right), \tag{5.4}$$

where $\tilde{S}(\varphi_i)$ is obtained from $\tilde{S}_P(U)$ by substituting $\exp(i\varphi_i)$ for U and $\exp(-i\varphi_i)$ for U^+. For $N \to \infty$ the expression further simplifies. In this limit the zero order approximation

of the steepest descent method becomes exact and (cf. e.g. Ref. [7])

$$F = \sum_{i=1}^{N} \tilde{S}(\bar{\varphi}_i) + \tfrac{1}{2} \sum_{i \neq j}^{N} \ln \sin^2 \left(\frac{\bar{\varphi}_i - \bar{\varphi}_j}{2} \right). \tag{5.5}$$

Here $\bar{\varphi}_1, \ldots, \bar{\varphi}_N$ are the real numbers, each from the interval $0 \leqslant \bar{\varphi}_i \leqslant 2\pi$, which when substituted for $\varphi_1, \ldots, \varphi_N$ maximize F. It is convenient to rescale the action by

$$S(\varphi) = \tilde{S}(\varphi)/N = \sum_{l=1}^{L} \frac{\beta_l}{l} \cos l\varphi \tag{5.6}$$

and to introduce for the numbers $\bar{\varphi}_1, \ldots, \bar{\varphi}_N$ a density distribution denoted $\varrho(\varphi)$ and normalized by

$$\int_0^{2\pi} \varrho(\varphi) d\varphi = 1. \tag{5.7}$$

By definition $\varrho(\varphi)$ should be non-negative

$$\varrho(\varphi) \geqslant 0. \tag{5.8}$$

Using $\varrho(\varphi)$ and $S(\varphi)$ one can rewrite relation (5.5) as

$$F = \int_0^{2\pi} S(\varphi)\varrho(\varphi) d\varphi + \tfrac{1}{2} \int_0^{2\pi} d\varphi \int_0^{2\pi} d\psi \ln \sin^2 \frac{\varphi - \psi}{2} \varrho(\varphi)\varrho(\psi). \tag{5.9}$$

The function $\varrho(\varphi)$ defines the vacuum structure i.e. the structure of the ground state of the system. It can be obtained by maximizing the functional F (5.9) under the constraints (5.7) and (5.8). The constraint (5.7) can be imposed by introducing a suitable Lagrange multiplier i.e. maximizing instead of F the functional

$$\tilde{F} = F + \mu \int_0^{2\pi} \varrho(\varphi) d\varphi. \tag{5.10}$$

6. Equations for the density $\varrho(\varphi)$

The maximum of the functional $\tilde{F}$ can occur either when the functional derivative with respect to $\varrho(\varphi)$ of the right hand side of relation (5.10) vanishes

$$S(\varphi) + \int_0^{2\pi} \varrho(\psi) \ln \sin^2 \frac{\varphi - \psi}{2} d\psi + \mu = 0, \tag{6.1}$$

or when

$$\varrho(\psi) = 0. \tag{6.2}$$

In general the range $0 \leqslant \varphi \leqslant 2\pi$ will split into sections $C_1, \ldots, C_M$, where equation (6.1) holds and into sections $\bar{C}_1, \ldots, \bar{C}_{M'}$, where $\varrho(\varphi) = 0$. Usually $M = M'$ except for the case when $M = 1$, $M' = 0$ and equation (6.1) holds in all the interval.

Equation (6.1) is not very convenient, because it contains the unknown parameter μ. Differentiating with respect to φ, we obtain on the sections C_i the equations derived by Gross and Witten in Ref. [7]:

$$\frac{\partial S}{\partial \varphi} + P \int_0^{2\pi} \varrho(\psi) \operatorname{ctg} \frac{\varphi - \psi}{2} \, d\psi = 0. \tag{6.3}$$

The P in front of the integral means that the principal value should be taken. It is important that a solution of equations (6.3) is not necessarily a solution of the original equation (6.1). Indeed, integrating equation (6.3) with respect to φ we find for φ on C_i

$$S(\varphi) + \int_0^{2\pi} \varrho(\psi) \ln \sin^2 \frac{\varphi - \psi}{2} \, d\psi = -\mu_i \qquad i = 1, \ldots, M. \tag{6.4}$$

There is nothing in the Gross-Witten equations to prevent the constants μ_i from being different from each other. Thus, in order to obtain solutions of equation (6.1) it is necessary to supplement the Gross-Witten equations by the subsidiary condition

$$\mu_1 = \mu_2 = \ldots = \mu_M. \tag{6.5}$$

Another constraint is that a solution of equation (6.1) is acceptable only if it satisfies the positivity condition (5.8).

For action (5.6) the Gross-Witten equations can be solved in closed form. Before presenting the solution, however, let us see how far one can go using simple thermodynamical arguments.

7. Thermodynamical discussion

It is convenient to interpret the range of φ as the circumference of the unit circle. Then all the solutions for $\varrho(\varphi)$ can be classified as follows (cf. Fig. 3). A solution belongs to class A_0, if equation (6.1) is satisfied in all the range of φ. It belongs to class $A_M(B_M)$, if there are M arcs where $\varrho(\varphi) = 0$, M arcs where $\varrho(\varphi)$ is a positive solution of equation (6.1) and $\varrho(0)$ is positive (vanishes). In the notation introduced in the preceding section, for class A_M the point $\varphi = 0$ belongs to one of the arcs C_i, while for class B_M it belongs to one of the arcs $\bar{C}_i$. In Fig. 3 only symmetric solutions with

$$\varrho(2\pi - \varphi) = \varrho(\varphi) \tag{7.1}$$

are drawn. This symmetry can indeed be proved [9].

In order to apply thermodynamics the following dictionary is helpful

β	inverse temperature $1/T$	(7.2)
$-F/\beta$	free energy	(7.3)
$-\int_0^{2\pi} S(\varphi)\varrho(\varphi)d\varphi/\beta$	internal energy	(7.4)
$-\tilde{F}/\beta$	potential Ω	(7.5)
$\frac{1}{2}\int\int d\varphi d\psi \varrho(\varphi)\varrho(\psi) \ln \sin^2 \frac{\varphi-\psi}{2}$	entropy	(7.6)
μ_i/β	chemical potential on arc C_i	(7.7)
type of solution	type of phase.	(7.8)

The system can be replaced by a gas of N non-interacting atoms free to move along the circumference of the unit circle. Since $N \to \infty$, the density $\varrho(\varphi)$ is used instead of a list of the individual coordinates φ_i.

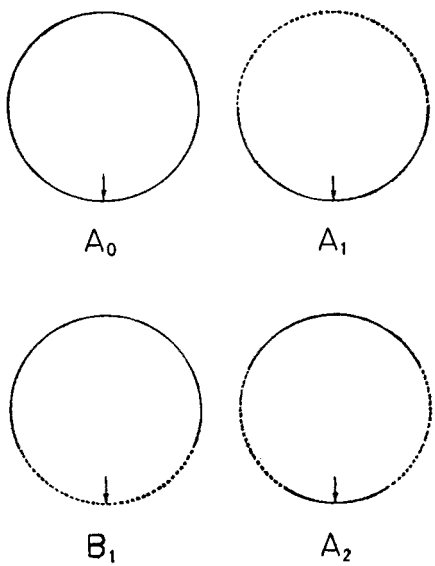

Fig. 3. Classification of phases. Continuous lines denote the arcs C_i, dotted lines denote the arcs $\bar{C}_i$. The point $\varphi = 0$ on each circle is indicated by an arrow

The constant β is some common scale of the constants $\beta_1, \ldots, \beta_L$ occurring in formula (5.6). A more precise definition will not be necessary. The result that $\tilde{F}[\varrho]$ should be maximized corresponds to the well-known theorem that at fixed temperature the equilibrium state is reached, when the free energy takes its minimum value. When besides the temperature the chemical potential is fixed instead of the number of particles, one has to minimize

instead of the free energy another thermodynamical potential denoted Ω. This corresponds to the maximization of $\tilde{F}[\varrho]$. In statistical physics at fixed temperature many parameters and functions can have sharply defined values only in the so called thermodynamical limit, when the number of degrees of freedom tends to infinity. Similarly in the present problem, the simple and dispersionless formulae for F and $\tilde{F}$ or $\varrho(\varphi)$ have been obtained due to the limit $N \to \infty$ [1].

At low temperature (weak coupling) the minimization of the free energy is almost equivalent to the minimization of the internal energy. Then all the gas concentrates near the highest maximum of $S(\varphi)$. In the most common case, when this occurs at $\varphi = 0$, the low temperature phase is A_1. According to the discussion from Sect. 3, therefore, the continuum is usually reached from the phase A_1 and all the other phases are just a nuisance.

When temperature increases, the entropy term gains in relative importance. Note that the entropy considered here has exactly the interpretation given to it by Boltzmann. It is the logarithm of the phase space available (in the space of all the unitary matrices U_P) when all the eigenvalues are fixed. The entropy term tends to increase the distances between the particles. Thus the arc C_1 centred on $\varphi = 0$ spreads. Now two things may happen. Either the arc spreads continuously untill it fills all the circumference and then there is a phase transition $A_1 \to A_0$, or, if there is another maximum of $S(\varphi)$,a new arc around this other maximum may form. For instance, if there is a maximum at $\varphi = \pi$, as is the case for some choices of the coupling constants when

$$S(\varphi) = \beta_1 \cos \varphi + \tfrac{1}{2} \beta_2 \cos 2\varphi, \tag{7.9}$$

the phase transition may be $A_1 \to A_2$.

In either case the name phase transition is correctly used. In thermodynamics there is a phase transition, when at some $\beta = \beta_0$ either there is a discontinuous change in the parameters of the system, or the evolution law changes with changing temperature. According to a classification due to Ehrenfest, when the k-th derivative of the free energy is discontinuous at $\beta = \beta_0$, while F itself and all its derivatives of order $n < k$ are continuous there, the transition at $\beta = \beta_0$ is of k-th order. When the number of arcs C_i changes it is almost obvious that the evolution law also changes. E.g. at the transition $A_1 \to A_0$ the arc C_1 stops spreading.

The transition $A_0 \leftrightarrow A_1$ is of third order as shown in Ref. [7]. It is therefore plausible that all the phase transitions, where an arc $\bar{C}_i$ disappears are of third order. On the other hand transitions, where a new arc C_i is formed at a finite distance from other arcs $C_{k \neq i}$ can be shown to be of lower order. This can of course be proved by inspection of the exact solutions, which we will derive. It is interesting, however, that a simple thermodynamic argument leads to the same conclusion. It is well known in thermodynamics that, if in a phase the system can be overheated, or overcooled, the corresponding transition is of first order. For instance for the transition $A_1 \leftrightarrow A_0$ it is not possible to imagine overcooling or overheating. On the other hand the phase transition $A_1 \to A_2$ can be delayed. It is possible to imagine that the arc C_1 will go on spreading and the arc C_2 will not be formed.

---

[1] The authors thank dr. J. Wosiek for a discussion concerning this point.

Formally this corresponds to the situation when both solutions A_1 and A_2 exist, but solution A_2 corresponds to a higher value of $\tilde{F}$. Calculations show that the corresponding phase transition is of second order in our model. The difference with respect to standard thermodynamics is understandable. When e.g. the temperature of water in equilibrium with its vapour is rised, all water goes over into vapour, however small is the change of β. Here small changes in β cause only flows of small quantities of the gas between C_1 and C_2. Consequently there is no discontinuity in the internal energy and no first order transition.

Also the question, when a new arc is formed has a rigorous answer. One easily shows (cf. Ref. [9]) that when $\mu_i > \mu_j$, particles will flow from arc C_i to arc C_j. The equilibrium condition is the one given by formula (6.5). Thus the chemical potentials defined by formula (6.4) have the properties known from standard thermodynamics. Bachas and Dashen [10] have recently suggested that a first order transition occurs, when a new maximum (in their paper minimum, because of a different convention) in the action $S(\varphi)$ appears. The present analysis gives instead a theorem rigorously valid for our model: The phase transition occurs, when the chemical potential of a particle at the new maximum equals the chemical potentials on the existing arcs C_i.

8. Solution for the strong coupling phase

For the strong coupling phase the Gross-Witten equations (6.3) are satisfied along all the circumference of the unit circle. In order to find the density $\varrho(\varphi)$ let us define on the complex plane the function

$$\Phi(z) = 2\int_C \frac{\tilde{\varrho}(\tau)d\tau}{\tau - z}. \tag{8.1}$$

The integration is over the closed contour $|\tau| = 1$. The function $\varrho(\tau)$ is related to the density $\varrho(\varphi)$ by

$$\tilde{\varrho}(e^{i\varphi}) = \varrho(\varphi) \quad \text{for} \quad e^{i\varphi} \quad \text{on} \quad C. \tag{8.2}$$

Formula (8.1) defines two holomorphic functions: the function $\Phi^+(z)$ holomorphic inside the unit circle and the function $\Phi^-(z)$ holomorphic outside. For $|z| = 1$ the function is undefined, but denoting by $\Phi^{\pm}(\tau)$ the limits of the functions $\Phi^{\pm}(z)$ for z tending to a point τ with $|\tau| = 1$, we have (cf. e.g. Ref. [11])

$$\Phi^+(\tau) - \Phi^-(\tau) = 4\pi i\tilde{\varrho}(\tau) \tag{8.3}$$

$$\Phi^+(\tau) + \Phi^-(\tau) = 2P\int_C \frac{\tilde{\varrho}(z)dz}{z - \tau} = \alpha(\tau) + 1. \tag{8.4}$$

The second equality in formula (8.4) follows from the Gross-Witten equations (6.3), because

$$\alpha(e^{i\varphi}) = \partial S/\partial\varphi. \tag{8.5}$$

It is seen from relations (8.3) and (8.4) that knowing the difference $\Phi^+(\tau)-\Phi^-(\tau)$ it is easy to calculate the sum $\Phi^+(\tau)+\Phi^-(\tau)$. In order to solve the inverse problem let define the function

$$\Psi^{\pm}(\tau) = \pm\Phi^{\pm}(\tau). \tag{8.6}$$

Substituting into the relations (8.3) and (8.4), we can use the known difference $\Psi^+(\tau)-\Psi^-(\tau)$ to derive

$$\Psi(z) = -\frac{i}{\pi}\int_C \frac{\alpha(\tau)+i}{\tau-z}\,dz \tag{8.7}$$

and

$$4\pi i\tilde{\varrho}(\tau) = -\frac{2i}{\pi}P\int_C \frac{\alpha(\tau)+i}{z-\tau}\,dz. \tag{8.8}$$

For $S(\varphi)$ given by formula (5.6) the integration can be explicitly performed and one finds

$$\varrho(\varphi) = \frac{1}{2\pi}\left[1+\sum_{l=1}^{L}\beta_l\cos l\varphi\right]. \tag{8.9}$$

It is seen that this density is correctly normalized to unity, but the positivity condition for all φ is satisfied only, if all the coefficients β_l are sufficiently small. Thus indeed this is a strong coupling phase.

For action (5.6) this phase always exists. It is enough, however, to perturb $S(\varphi)$ in the vicinity of one point $\varphi = \varphi_0$ so that

$$\lim_{\varepsilon\to 0}\left[(\partial S/\partial\varphi)_{\varphi=\varphi_0-\varepsilon}-(\partial S/\partial\varphi)_{\varphi=\varphi_0+\varepsilon}\right] > 0 \tag{8.10}$$

and the integral (8.8) for $z = \exp(i\varphi_0)$ will diverge to minus infinity. Thus the positivity condition cannot be satisfied and the strong coupling phase does not exist. This explains the absence of the phase transition for Manton's action (cf. Ref. [8]). It also shows how unstable the phase structure is with respect to changes in the single plaquette action.

9. *Solutions for other phases*

Also in the case of phases other than A_0 the Gross-Witten equations can be solved by standard methods (cf. e.g. Ref. [11]). The definition (8.1) remains valid, except that now C denotes the sum of arcs, where the Gross-Witten equations are satisfied. Equation (8.2) holds only on C, while on $\bar{C}$ the function $\tilde{\varrho}(\tau)$ is unknown. Equation (8.4) is replaced by the pair of equations

$$\Phi^+(\tau)+\Phi^-(\tau) = \alpha(\tau)+i \quad \text{on} \quad C \tag{9.1}$$

$$\Phi^+(\tau)-\Phi^-(\tau) = 0 \quad \text{on} \quad \bar{C}. \tag{9.2}$$

Thus neither the sum, nor the difference of the functions Φ^+ and Φ^- is known on all the curve $|\tau| = 1$. In order to circumvent this difficulty, one introduces the auxiliary function

$$R_M(z) = \sqrt{\prod_{j=1}^{2M} (z - \exp(i\psi_j))}, \tag{9.3}$$

where $\exp(i\psi_j)$ for $j = 1, \ldots, 2M$ are the end points of the arcs C_i. The cuts of the square root go along the arcs C_i, so that labeling as before by superscripts $\pm$ the functions inside and outside the unit circle, we have for z tending to a point τ on the circumference

$$R_M^+(\tau) = \begin{cases} -R_M^-(\tau) & \text{on} \quad C \\ +R_M^-(\tau) & \text{on} \quad \bar{C}. \end{cases} \tag{9.4}$$

Therefore

$$\left[\frac{\Phi(\tau)}{R_M(\tau)}\right]^+ - \left[\frac{\Phi(\tau)}{R_M(\tau)}\right]^- = \begin{cases} \dfrac{\alpha(\tau)+i}{R_M^+(\tau)} & \text{on} \quad C \\ 0 & \text{on} \quad \bar{C} \end{cases} \tag{9.5}$$

and proceeding as in the previous section

$$\tilde{\varrho}(\tau) = \frac{-R_M^+(\tau)}{2\pi^2} \int_C \frac{\alpha(z)+i}{R_M^+(z)(z-\tau)}\, dz. \tag{9.6}$$

Substituting the solutions (8.8) and (9.6) back into the Gross-Witten equation, one finds that $\tilde{\varrho}(\tau)$ given by formula (8.8) always is a solution, while $\tilde{\varrho}(\tau)$ given by formula (9.6) is a solution only if the M subsidiary conditions

$$\int_C \frac{\alpha(\tau)+i}{R_M^+(\tau)} \tau^k d\tau = 0; \quad k = 1, \ldots, M-1 \tag{9.7}$$

are satisfied. For given C these are constraints on the action $S(\varphi)$. In our case $S(\varphi)$ is given, and the constraints (9.7) together with the constraints (6.5) limit the possible choices of M and of the end points $\psi_j, j = 1, \ldots, 2M$. It follows from the general theory (cf. Ref. [11]) that, when conditions (9.7) are not satisfied, the Gross-Witten equations have no bounded solutions. Solutions unbounded, but integrable, exist in general and cannot be discarded a priori. One can show, however, (cf. Ref. [9]) that they cannot correspond to maxima of the functional $\tilde{F}$ and consequently are of no interest for our work.

The density $\varrho(\varphi)$ is given by the formula

$$\varrho(\varphi) = \begin{cases} \tilde{\varrho}(e^{i\varphi}), & e^{i\varphi} \quad \text{on} \quad C \\ 0, & e^{i\varphi} \quad \text{on} \quad \bar{C}. \end{cases} \tag{9.8}$$

The function $\tilde{\varrho}(\varphi)$ for $\exp(i\varphi)$ on $\bar{C}$ is also of interest, however, because the constraints (9.7) can be replaced (cf. Ref. [9]) by the equivalent and more convenient set of constraints

$$\int_{\bar{C}_i} \tilde{\varrho}(\varphi) d\varphi = 0; \quad i = 1, \ldots, M. \tag{9.9}$$

This completes the analytic solution of our problem. Integral (9.6) can be explicitly evaluated (cf. Ref. [9]), but the formula is not very instructive and we do not quote it here. The time consuming part of a practical calculation for given coupling constants, i.e. for given $\beta_1, \ldots, \beta_L$, is to find using conditions (6.5) and (9.9), which values of M correspond to acceptable ($\varrho \geqslant 0$) solutions and what are the corresponding end points $\psi_1, \ldots, \psi_{2M}$. There is a general theorem (cf. Ref. [9]) that phases with $M > L$ are impossible. This reduces somewhat the necessary work.

For $L = 2$ it is possible to perform all the calculations analytically and one obtains the phase diagram shown in Fig. 4.

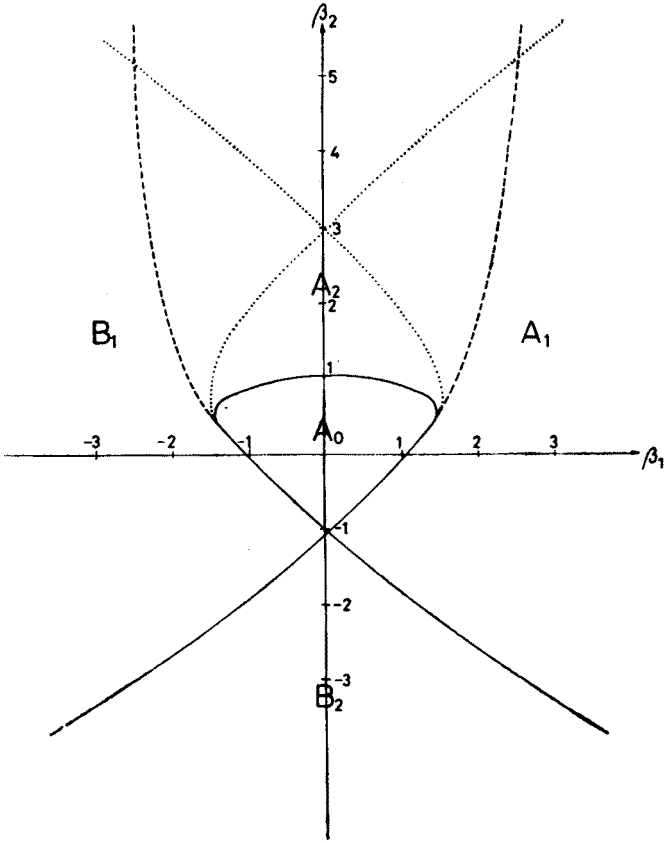

Fig. 4. Phase diagram for action (5.6) with $L = 2$. The continuous lines denote the third order transitions, the dashed lines denote the second order transitions. The dotted lines denote the existence limits for the metastable phase A_1 and B_1

10. Conclusions

Our work shows that even in a highly simplified model the phase structure may be arbitrarily rich and is sensitive to details in the functional form of the single plaquette action.

Both pessimistic and optimistic conclusions can be drawn from this result. The bad news is that there is no chance to explore a significant fraction of possibilities by choosing simple actions, performing Monte-Carlo calculations and publishing the phase diagrams obtained.

On the other hand, since the phase structure is so flexible and potentially so rich, there is a good chance to find an action corresponding to a favourable phase diagram, whatever other requirements are imposed on the action. For instance optimizing the convergence to the continuum limit [12], or trying to improve the signal to noise ratio in the search for nonperturbative effects [13] should not be connected to some unavoidable deterioration of the phase structure.

One of the authors (KZ) thanks the organizers of the Smolenice Conference for their hospitality.

REFERENCES

[1] K. Johnson, *Acta Phys. Pol.* **B6**, 865 (1975).
[2] A. Hasenfratz, P. Hasenfratz, *Phys. Lett.* **93B**, 165 (1980).
[3] R. G. Dashen, D. J. Gross. *Phys. Rev.* **D23**, 2341 (1981).
[4] A. Gonzalez Arroyo, C. P. Korthals Altes, J. Peiro, M. Perrottet, *More structure in the string tension*, preprint (1982).
[5] K. Wilson, *Phys. Rev.* **D10**, 2445 (1974).
[6] N. S. Manton, *Phys. Lett.* **96B**, 328 (1980).
[7] D. J. Gross, E. Witten, *Phys. Rev.* **D21**, 446 (1980).
[8] C. Lang, P. Salomonson, B. S. Skagerstam, *Nucl. Phys.* **B190** [FS3], 337 (1981).
[9] J. Jurkiewicz, K. Zalewski, *Vacuum structure of the $U(N \to \infty)$ gauge theory on a two dimensional lattice for a broad class of variant actions.* Marseille preprint CPT-82/P. 1438 (1982).
[10] R. Bachas, R. Dashen, *Phase transitions and local minima of the action*, Princeton preprint (1982).
[11] N. I. Muschelishwili, *Singular Integral Equations* (in Russian), second edition, Moscow 1962.
[12] K. Symanzik, *Some topics in quantum field theory*, DESY preprint 81-068 (1981).
[13] A. Di Giacomo, G. Paffuti, *Nucl. Phys.* **B205** [FS5], 313 (1982).