Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

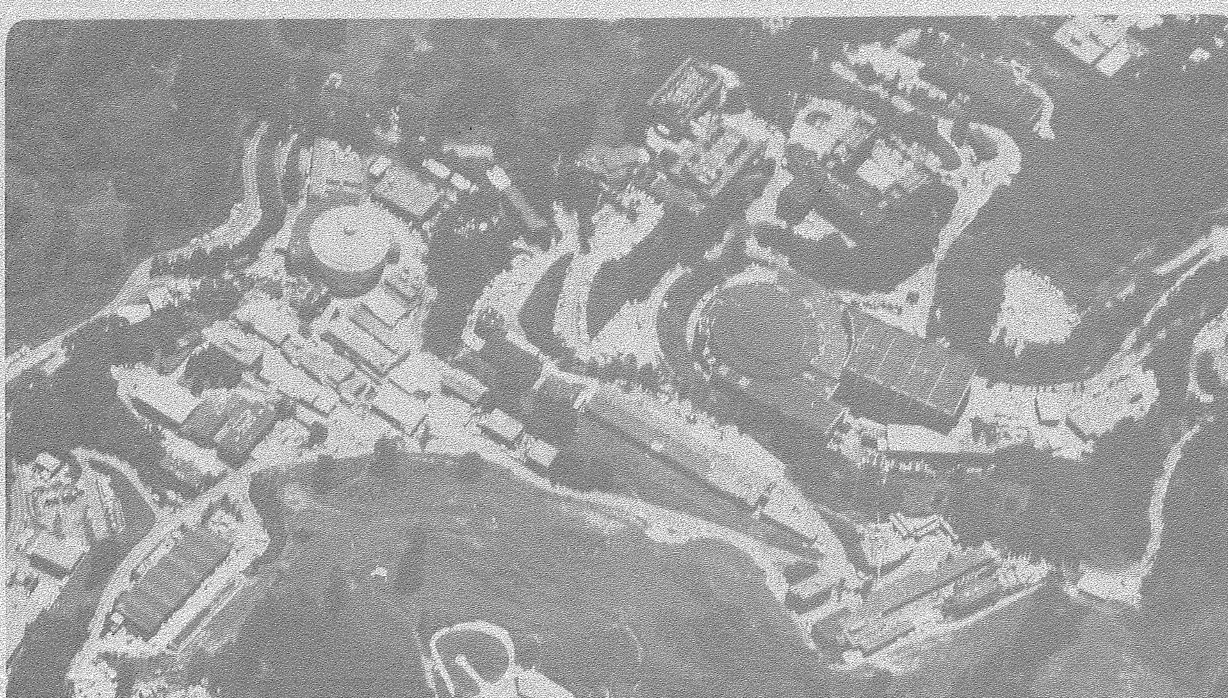
Physics, Computer Science & Mathematics Division

Submitted to Nuclear Physics B

DOMINANT EUCLIDEAN CONFIGURATIONS FOR ALL N

K. Bardakci, D. G. Caldi, and H. Neuberger

April 1980



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

LBL-10885
UCB-PTH-80/2 April 1980

DOMINANT EUCLIDEAN CONFIGURATIONS
FOR ALL N

K. Bardakci *†, D. G. Caldi † and H. Neuberger *

Department of Physics

and

Lawrence Berkeley Laboratory

University of California

Berkeley, California 94720

* Research supported in part by the National Science Foundation under Grant No. PHY-79-23251.

† Research supported in part by the High Energy Physics Division of the U. S. Dept. of Energy under contract No. W-7405-ENG-48.

Abstract

We identify a class of Euclidean configurations which appear to be dominant in the functional integral of the CP^{N-1} models. These configurations are point-like topological excitations, and they may be viewed as constituents of instantons, although they are defined independently of instantons through a continuum duality transformation. We show not only that these configurations survive as $N \to \infty$, but that in the plasma phase they are responsible for the effects encountered within the $1/N$ expansion — confinement, θ-dependence, and dynamical mass generation. We also discuss the possible types of plasma phase for a statistical mechanical model of the interacting excitations. Possible generalization to QCD is briefly discussed.

I. INTRODUCTION

The usefulness and importance of semiclassical methods in the study of quantum field theory have been amply demonstrated over the past few years[1]. However, the role of semiclassical approximations in the Euclidean functional integral of a scale-invariant quantum field theory[2] has become a matter of some disagreement of late. The disagreement centers around two related questions. The first is that of adequately identifying those field configurations which are statistically important in the partition function or functional integral so that a sensible calculation can, at least in principle, be carried out. The second question concerns another approximation scheme for quantum field theory, the $1/N$ expansion[3,4], and its relation to the semiclassical approximation. It has recently been argued[5,6] that the $1/N$ expansion is the much preferred method of attack, as suggested in two-dimensional theories, and that the effects calculated by semiclassical approximations, at least around the best known classical field configurations, instantons, are negligible. In this paper we shall study these questions with a view to resolving them by exhibiting field configurations which appear to be important in the functional integral, and furthermore survive in the large N-limit to the extent that they may even be viewed as driving the results obtained there. Our study here will be concerned with certain two-dimensional models, the recently discovered CP^{N-1} models, where the issues are more clearly drawn and the complexities, though sufficiently difficult as often to preclude exact computation, nevertheless do not obscure the essentials of the physics. The situation for four-dimensional quantumchromodynamics (QCD) is much more problematical, from both the 1/N point of view as well as the semiclassical, so we shall have to limit ourselves in this paper to some more general comments and plans of attack.

Let us begin by defining the models, albeit briefly since a fairly extensive literature[5-17] on them already exists. The CP^{N-1} model is the theory of minimally coupled complex rays in N-dimensional complex space. The action may be written as

$$\left.\begin{aligned}
& S = \frac{1}{g^2} \int d^2x \, (\mathcal{D}_\mu z_\alpha)^* \, (\mathcal{D}_\mu z_\alpha), \\
& \text{with the constraint} \\
& z^*_\alpha z_\alpha = 1, \qquad\qquad \alpha = 1, \ldots N \\
& \text{where} \\
& \mathcal{D}_\mu = \partial_\mu + iA_\mu, \\
& \text{and} \\
& A_\mu = iz^*_\alpha \partial_\mu z_\alpha
\end{aligned}\right\} \tag{1.1}$$

by its equations of motion. Hence, besides a global SU(N) symmetry, the model possesses a local U(1) gauge invariance. This and the constraint results in the fields z_α taking values in N-1 dimensional complex projective space CP^{N-1}: the space of equivalence classes of complex N-vectors under complex scalar multiplication. For N = 2 the model is equivalent to the familiar O(3) non-linear sigma model.[18-25].

In two dimensions the models are asymptotically free, conformally invariant, and topologically non-trivial—features which they share with QCD, thus making them presumably good laboratories for exploring the possible consequences of these features for QCD. (It may be of interest to note that asymptotic freedom follows simply because the theory is one of minimally coupled fields which take values in a compact, homogeneous manifold [26]).The model is also 1/N expandable since the theory may be easily rewritten in terms of $SU(N)_{flavor}$-invariant quantities. The major result [5,10] of the 1/N expansion is that the model may be viewed as describing N charged particles, with a dynamically generated mass, which interact at large distances through Coulomb forces. This leads to the disappearance of charged particles from the spectrum and to the physical significance of the θ-parameter. Furthermore, after the introduction of fermions into the model, the chiral $U(1)_A$ problem is resolved. Fortunately, the spectrum, indeed the whole S-matrix, of the model for N = 2 is exactly known [25].(Presumably a set of eigenstates could now be built.) So the results of the 1/N expansion have been shown to be qualitatively correct. Some authors[10,11,12,27] have emphasized the importance of topologically non-trivial field configurations in obtaining these results even in the 1/N expansion. Others[5,6],however, have chosen to stress the view of the 1/N expansion as a summation of perturbation theory Feynman diagrams.

Topologically non-trivial configurations are known to exist in the CP^{N-1} models since they have a local U(1) gauge symmetry, while on the classical level they exhibit spontaneous breaking of the

SU(N) global symmetry. It follows that they have instanton solutions to the Euclidean equations of motion for all N. Such configurations are able, in principle, to produce dramatic effects in two dimensions (e.g. confinement in spite of the Higgs mechanism[27]) of exactly the same type as does the 1/N expansion.

Most calculations with instantons have, until recently, been done in the dilute gas approximation, which is infrared divergent in scale-invariant theories like QCD or CP^{N-1}, and so requires a cut-off. Lately, a number of calculations[12,13,23,29,30] have appeared using multi-instanton solutions instead. Leaving aside the results for QCD, which are still incomplete, the results in CP^{N-1} appear to indicate that the gas of instantons is dense and infrared finite, and hence that the dilute gas configurations are statistically negligible. However, it should be noted that the exact multi-instanton solutions are a very restrictive set of configurations, and so almost certainly they also are irrelevant statistically. Furthermore, incoherently summing the contributions from pure instantons and pure anti-instantons does not appear to satisfy the cluster decomposition theorem[24,30]. Physically, the problems with cluster decomposition arise because multi-instanton configurations are very strongly correlated. Making a specific choice for the field at some point dictates the choice to be made at all other points, so long as one wants an exact solution to the equations of motion. The apparent way to overcome the cluster decomposition problem is to include instanton - anti-instanton configurations somehow. This would also give a statistically more significant set. Indirectly this is what we propose to do here. These difficulties, however, do not render useless the impressive calculations done by saturating the functional integral with only multi-instanton (multi-anti-instanton) solutions and the small fluctuations around them. Besides the significant information these calculations provide about what happens to the instanton gas, they also give quite reasonable results[23,24] for some Green's functions, where the cluster decomposition problem does not occur.

But in general one must sum coherently both instantons and anti-instantons together, and this is what the dilute gas approximation does. So long, then, as one considers only small scales, the dilute gas is relevant. Furthermore, and perhaps more compelling, in QCD it is the only known tractable approximation. One can only hope that in trying to extend the dilute gas approximation as far as possible, one may be able to obtain some signals about what happens beyond its range of validity. There are some indications [31] that this may be the situation in QCD.

However, as already noted, the whole instanton approach has recently been questioned in the context of the 1/N expansion. The principal argument[5,15,17,32,33] is that whatever the instanton effects may be, they become irrelevant as $N \to \infty$ since they vanish like e^{-cN}. Since the 1/N expansion is smooth and reliable, this argument says that instanton effects may be negligible even for $N = 2$. There are strong indications[4] that the $N \to \infty$ limit of QCD exhibits many of the rough features of hadronic physics, and thus the leading terms in the 1/N expansion may be a good approximation for $N = 3$. Thus, the argument continues, as in CP^1, the instanton contribution to such effects as confinement and the resolution of the $U(1)_A$ problem may be irrelevant. These considerations are especially directed against the dilute gas approximation. But even there it could still turn out that, due to higher order effects (including, perhaps, interactions), the large-N behavior is not that given by the one-loop calculation.

Other, and in our opinion, weaker objections to instantons have been advanced[5].The exact instanton (or multi-instanton) solutions are obtained by imposing boundary conditions at infinity which ensure finite action. The boundary conditions at infinity should be those which are obeyed by the most important vacuum configurations (we use Euclidean language). In two dimensions because of kinematics, and in four dimensions because of confinement viewed as an experimental fact, these are certainly not finite action boundary conditions. But it is almost certain that, as opposed to real time, in imaginary time exact solutions of the equations of motion, or other finite action configurations are not important. This is so even in the dilute gas case. But what goes against the dilute gas is that from both the 1/N picture, as well as the multi-instanton calculations, it appears that one should envisage the appearance of a nonzero density of topological charge as a result of the contribution not of small lumps of concentrated topological charge density, but rather of a quite smooth distribution.

Another argument[5,6] against instantons is that the effective action derived to perform the 1/N expansion does not possess any instanton-like solutions. Hence, given the reliability of the 1/N expansion, instantons cannot be of consequence to the quantum theory. However, it appears that this argument may be too facile, since it has since been shown[27] that instantons do appear in a different form, namely as poles of the effective 1/N integrand, and hence their contribution may be evaluated by closing the contour in a way different from the 1/N expansion. So not only do instantons survive but they may, and at least in some models[27] do, give the same results as the 1/N expansion.

It seems to us that for the present it is far from obvious which approach to long-distance QCD is potentially more rewarding -the 1/N expansion or the quasiclassical approximation. The qualitative success of 1/N should not have a destructive influence on the quasiclassical approach. Our purpose in the present paper is to set up a framework for the quasiclassical approach to the CP^{N-1} models. We require our configurations to contain the dilute gas case as a subset, which presumably is insignificant. We also will present arguments to show that our configurations do not disappear in the large N-limit, nor do they predict a distribution of dilute lumps of topological charge density. Furthermore, the configurations are intimately related to instantons, yet, in the plasma phase, they produce effects quite similar to those obtained in the 1/N expansion, that is, confinement and the related θ-dependence, as well as dynamical mass generation. Hence our work supports the point of view of those who stressed the importance of topologically non-trivial configurations even in obtaining the results of the 1/N expansion.

The plan of the rest of the paper is as follows. In section II we treat only the case N=2 (essentially the O(3) non-linear sigma model). We present the configurations, which we will call interchangeably topological excitations, vortices, instanton quarks, or merons, and which are essentially equivalent to the merons of ref.[20]. Here we also set up a field theoretical framework, via what is known as the duality transformation[34] or functional Fourier transformation [35], to take their contribution into account. In section III the derivations for arbitrary N are presented. We argue, in section IV, in favor of the importance of the configurations for confinement and the 1/N expansion. Section V presents a model for the treatment of the interactions, especially the spin-spin part, of the vortices. Section VI concludes the paper with a summary and a brief discussion of probable implications for QCD. In an appendix, some technical analysis related to section V is presented.

II. TOPOLOGICAL EXCITATIONS IN CP^1

The dilute gas picture rests on configurations made out of single, well separated instantons and anti-instantons. This implies a parametrization naturally attached to the one instanton solutions. Such a parametrization views an instanton as an elementary excitation having a variable size, a center, and some angular parameters which determine its orientation in the space of a compact group. Such a parametrization is not suited for the description of a larger set of configurations than the dilute set. The first step in finding a larger set of configurations is, therefore, to find a new parametrization of the dilute gas, the relevance of which is not dependent on diluteness. A good hint of what that parametrization might be is given by the multi-instanton solutions. Although these configurations are extremely correlated and, therefore, presumably not statistically important, they do point out the existence of an easily generalizable parametrization.

The one instanton solution may be written in two equivalent ways: in the standard parametrization,

$$z_\alpha^{(1)} = \frac{\lambda u_\alpha + [(x_1 - x_1^o) - i(x_2 - x_2^o)]\, v_\alpha}{[\lambda^2 + (x - x^o)^2]^{\frac{1}{2}}}, \tag{2.1}$$

$$\alpha = 1,2$$

with $\quad u_\alpha^* u_\alpha = v_\alpha^* v_\alpha = 1, \quad u_\alpha^* v_\alpha = 0;$

and in a different one,

$$w_{(1)} \equiv \frac{z_1}{z_2} = c\,\frac{s-a}{s-b}\ , \quad s = x_1 - ix_2\ , \ a,b,c \in \mathbb{C}\ . \tag{2.2}$$

The relation between the two parametrizations is nonlinear for $c \neq 1$. Thus they are genuinely different. Eq.(2.2) has mixed some of the space and group degrees of freedom of Eq. (2.1). It is Eq. (2.2) which immediately generalizes to the exact n-instanton solutions:

$$w_{(n)} = c\,\frac{\prod_{i=1}^{n} (s - a_i)}{\prod_{i=1}^{n} (s - b_i)}\ . \tag{2.3}$$

Thus we are led to the conclusion that we should think in terms of some local point-like excitations which come in pairs to make instantons. The parameters would be the locations of these point-like objects. That this parametrization of the exact solutions is adequate is shown by the computation of small fluctuations around them. It has been shown in two calculations [12,23] that the contribution to the partition function of the Gaussian fluctuations around an n-instanton solution is given by

$$\begin{aligned}\mathcal{Z}^{\text{inst}}_{(n)} = \frac{M^{2n}}{(n!)^2} \int d^2a_1 \ldots d^2a_n d^2b_1 \ldots d^2b_n \frac{d^2c}{(1+|c|^2)^2} \times \\ \exp\left\{ - \left[\sum_{i,j} \ln |a_i - b_j|^2 - \right.\right. \\ \left.\left. \sum_{i<j} (\ln|a_i - a_j|^2 + \ln|b_i - b_j|^2) \right] \right\}\ .\end{aligned} \tag{2.4}$$

We see that the contribution of c factorizes and that the a's and b's play the role of the locations of positive and negative charges interacting through Coulomb forces in two-dimensional space.

The interpretation of the a's and b's as locations of charged particles is supported by the classical dynamics of the system. Indeed, as shown in refs. [20,22] , an instanton interacts classically with a far away anti-instanton (both scales are small compared to the distance between the centers) by a dipole-dipole type interaction. It turns out that each (anti)instanton may be associated with an electric dipole having its constituent charges situated precisely at a and b. We should be cautious, however; instantons do not interact with themselves classically [20,22]. Thus it appears necessary to think about an additional parameter, some kind of discrete spin which takes care of this fact [22].

We want now to look for a characterization of the positions a and b which frees them from the direct relation to exact solutions of the equations of motion. (At this point we differ conceptually somewhat from ref. [20].) Eq.(2.2) suggests that a gauge invariant characterization of a and b is given by defining points in Euclidean space around which the phase of w rotates. This characterization is very suggestive because it is essentially the definition of vortex configurations in two dimensions, and it is known that in any reasonable theory [36] vortices will tend to interact logarithmically, that is by Coulomb forces. A vortex has an ultraviolet-divergent self-energy and therefore one expects the dynamics to smear out the singularity. This is achieved by making w either vanish or blow up at the locations of the vortices. The existence of these two possibilities is a reflection of a trivial discrete symmetry of Eq. (1.1) for N = 2, that is $z_1 \leftrightarrow z_2^*$, or, in terms of the gauge invariant unconstrained field w, $w \leftrightarrow \frac{1}{w^*}$. This discrete symmetry leaves invariant the nature of the vortex, and it leads us to the following parametrization of the z and w fields:

$$z_1 = f(\rho_1)\, e^{i(\theta_+ + \frac{1}{2}\phi)} \qquad 1 \geq f \geq 0$$

$$z_2 = \sqrt{1 - f^2(\rho_2)}\; e^{i(\theta_+ - \frac{1}{2}\phi)} \qquad -\frac{\pi}{2} < \theta_+ < \frac{\pi}{2} \qquad (2.5)$$

$$-\pi < \phi < \pi$$

f being an undetermined function. In these new variables the partition function of the CP^1 model is given by

$$\mathcal{Z} = \int \prod_x [d^2 z_1 d^2 z_2]\, e^{-\frac{1}{g^2} S[z]} \prod_x [\delta(|z|^2 - 1)]$$

$$= \mathcal{N} \int \prod_x [d\rho_1\, d\rho_2\, d\theta_+\, d\phi] \prod_x \left[\frac{d}{d\rho} f^2\right] \prod_x [\delta(\rho_1 - \rho_2)] \times$$

$$e^{-\frac{1}{g^2} S(\rho_1, \phi)} \quad , \qquad (2.6)$$

with

$$S(\rho, \phi) = \int \left[f^2(1 - f^2)\, (\partial_\mu \phi)^2 + \left(\frac{df}{d\rho}\right)^2 \frac{(\partial_\mu \rho)^2}{1 - f^2} \right] , \qquad (2.7)$$

or equivalently

$$S(\rho, \phi) = \int \left[f^2(1 - f^2)(\partial_\mu \phi)^2 + \left(\frac{1}{f}\, \partial_\mu \sqrt{1 - f^2}\right)^2 \right] . \qquad (2.8)$$

We see that the dependence on θ_+ has disappeared and this may be an alarming sign: didn't we throw away singular gauge configurations which are important and have nonvanishing action? The answer is that we did not, and this point deserves some further comments.

Since there is no classical kinetic energy term for the gauge field in the action, singular gauge transformations do not cost a finite action. This means that, in the (2.1) parametrization, zero size instantons have zero action. In the picture of vortices this is obvious: opposite vortices annihilate when located at the same point. But this contradicts the standard derivation of self-duality equations, which proceeds by obtaining lower bounds for the action in each topological sector. Indeed, a careful examination of the derivation shows that it goes wrong, as it should, for zero size instantons. One uses the identity

$$(\mathcal{D}_\mu z_\alpha)^* \mathcal{D}_\mu z_\alpha = \tfrac{1}{2} \sum_\alpha |\mathcal{D}_\mu z_\alpha \pm i\varepsilon_{\mu\nu} \mathcal{D}_\nu z_\alpha|^2$$

$$\mp i\varepsilon_{\mu\nu}\partial_\mu z^*_\alpha \partial_\nu z_\alpha . \tag{2.9}$$

This identity holds for regular or singular A_μ's. The next step is to write

$$\varepsilon_{\mu\nu}\partial_\mu z^*_\alpha \partial_\nu z_\alpha = \varepsilon_{\mu\nu}\partial_\mu (z^*_\alpha \partial_\nu z_\alpha) = i\varepsilon_{\mu\nu}\partial_\mu A_\nu . \tag{2.10}$$

This is wrong for a zero size instanton since at the point where the singularity sits $\varepsilon_{\mu\nu}\partial_\mu\partial_\nu z_\alpha \neq 0$. This point indicates that the definition of topological charge by the line integral $\frac{1}{2\pi}\oint A_\mu dx_\mu$ is not always appropriate, whereas the local definition via vorticity avoids this problem. We should remark here that the fact that the action of an n-instanton configuration undergoes a discontinuous jump when an "a" touches a "b", is important in the interpretation of the result in Eq.

(2.4) obtained in refs. [12, 23]. In general one would expect that entropy will be an important enough criterion to make the configurations which have some "a" sitting on top of a "b" irrelevant statistically. Unfortunately, at the specific point of Eq. (2.4) the interaction just takes over, and for a Coulomb gas of point-like objects we are at a phase transition, from a plasma phase to a collapsing dipole phase. In ref.[23] the additional ultraviolet divergences which appear have been treated by the introduction of a cut-off, in which case the system is above the phase transition and in the plasma phase. When an "a" hits a "b", the classical action changes discontinuously. This supports the introduction of an additional cut-off since the limit $a \to b$ need not be reliable.*

The possible presence of vortices in Eq. (2.8) is signaled by the

---

* Eq. (2.4) is not free of ultraviolet divergences, despite standard one-loop renormalization which generated the mass M. In the field theoretical context this additional divergence has been observed by Lehmann and Stehr [37] and by Schroer and Truong [38] who showed that when the sine-Gordon system approaches the point where solitons unbind from their parent mesons (in the fermion, massive Thirring model this is where the current-current coupling goes to zero); the normal ordering employed in the boson theory no longer takes care of all the u.v. divergences. Eq. (2.4) represents the perturbation theory in the cosine term of the sine-Gordon Lagrangean with only normal-ordering regularization [39].

fact that ϕ as an angular variable need not satisfy the equation $\varepsilon_{\mu\nu}\partial_\mu\partial_\nu\phi = 0$ everywhere. The reason is that $\oint dx_\mu \partial_\mu \phi$ does not have to be zero around every point in space; it may be an integer multiple of 2π . Let us denote for convenience

$$\Theta_\mu = \varepsilon_{\mu\nu}\partial_\nu\phi \; . \tag{2.11}$$

One may now consider $f^2(1-f^2)$ as being related to an effective dielectric constant, ε:

$$\varepsilon = \frac{1}{4f^2(1-f^2)} \quad ; \quad \varepsilon \geqslant 1 \; . \tag{2.12}$$

The second term in the action (2.8) defines a distribution for the dielectric constant of the medium. Θ_μ is somewhat similar to the "D" field, the electric field produced by the "external" charges only. But Θ_μ should not be considered only an electric field. This is so because curl $\vec{E}$ = curl $\frac{1}{\varepsilon}\vec{D}$ must not vanish. So the system also contains a current distribution (perpendicular to our 2-dimensional world) which is equivalent to a smooth background of magnetic charge. It will turn out that this smooth background can be formally integrated out. The divergence of Θ_μ can have contributions only from isolated quantized singularities,

$$\partial_\mu\Theta_\mu = 2\pi \sum_{j=1}^{m} q_j \;\; \delta(x-x_j) \; ; \; q_j \varepsilon Z \; . \tag{2.13}$$

By inspection it is clear that in order to avoid ultraviolet singularities one must have $f = 0$ or 1 at each x_j. Assuming that f is not 0 or 1 uniformly at infinity we find that, in order to avoid an infrared divergence, the following neutrality condition must be met,

$$\sum_{j=1}^{m} q_j = 0 \tag{2.14}$$

The topological charge $\hat{Q}$ is given by

$$\hat{Q} = \frac{1}{2\pi} \int d^2x \Theta_\mu \partial_\mu (f^2)$$

$$= \frac{1}{2\pi} \int d^2x \partial_\mu (\Theta_\mu f^2) - \sum q_j f^2(x_j) . \tag{2.15}$$

The first term is zero since f^2 is bounded and the system is neutral. Thus

$$\hat{Q} = - \Sigma' q_j , \tag{2.16}$$

where the prime means exclusion of those terms for which $f^2(x_j) = 0$. Clearly, with a given distribution of singularities one achieves the maximum of $|\hat{Q}|$ by requiring $f^2(x_j) = +1$ at all positive electric charges and $f^2(x_j) = 0$ at all negative electric charges (or vice versa). Since the action satisfies

$$S = \int d^2x \left\{ [\frac{1}{f} \partial_\mu (1 - f^2)^{\frac{1}{2}} \pm f(1-f^2)^{\frac{1}{2}} \Theta_\mu]^2 \right\} \pm 2\pi\hat{Q}$$

$$\Rightarrow |S| \geqslant 2\pi |\hat{Q}| , \tag{2.17}$$

a solution to the equations of motion may be achieved by looking for functions which satisfy the self-duality conditions:

$$\frac{1}{f}\,\partial_\mu (1-f^2)^{\frac{1}{2}} = \mp\, f(1-f^2)^{\frac{1}{2}}\,\Theta_\mu$$

$$\Rightarrow \frac{1}{2}\partial_\mu \log\left(\frac{1}{f^2} - 1\right) = \mp\,\Theta_\mu$$

$$\Rightarrow \frac{1}{2}\,\nabla^2 \log\left(\frac{1}{f^2} - 1\right) = \mp\, 2\pi \sum_{j=1}^{m} q_j\, \delta(x - x_j) \tag{2.18}$$

$$\Rightarrow \log\left(\frac{1}{f^2} - 1\right) = \mp \sum_j q_j \log (x - x_j)^2 + \Lambda$$

$$\text{with} \quad \nabla^2\Lambda = 0 \quad .$$

Now demanding $f^2 \underset{|\vec{x}|\to\infty}{\not\longrightarrow} 1$ or 0 forces Λ = const. Thus we obtain

$$\log \frac{1}{f^2} - 1 \underset{x\to x_j}{\sim} \mp\, q_j \log (x - x_j)^2$$

$$\Rightarrow f^2(x_j) = \begin{cases} 1 & q_j < 0 \\ 0 & q_j > 0 \end{cases} \Bigg\} \begin{array}{l}\text{for upper sign} \\ \text{(instantons)}\end{array} \qquad \begin{cases} 1 & q_j > 0 \\ 0 & q_j < 0 \end{cases} \Bigg\} \begin{array}{l}\text{for lower sign} \\ \text{(anti-instantons)}\end{array} \quad . \tag{2.19}$$

We see that self-duality forced $|\hat{Q}|$ to be maximal. Were this a result of self-duality alone, we might hope that enforcing a different distribution of the values 0 and 1 of f at the locations of the electrical charges, would lead to a well-defined instanton-anti-instanton configuration. Unfortunately there are no non-self-dual, finite action, regular solutions to the equations of motion of the O(3) sigma model [40].

In principle, a possible approach to the problem of finding the statistically relevant contributions within a given set of configurations has been given in ref. [41]. In practice, however, one is faced with a difficult technical problem since one needs to find a set of constraints, which depend on arbitrary continuous parameters, such that by varying the parameters one spans the set of configurations. Of course, the classical constraint equations must be exactly soluble for any value of the parameters. It is hard to see how this method may lead to tractable calculations for the case of instantons and anti-instantons since one needs to invent a constraint which, when a <u>continuous</u> parameter is varied, turns instantons into anti-instantons. Thus we shall resort to a different, more formal and less powerful method.

In order to single out the contributions of the vortex configurations to the partition function we should constrain the functional integration to a well-defined configuration of vortices and sum over all the possible configurations. Stating a well-defined constraint is physically more meaningful than stating what the complete field configuration is. This point has been explained and emphasized in ref. [41]. Furthermore, we should point out that our constraint is point-like, and so allows for the possibility of deriving a local field theory for the new excitations. This can be done by formal manipulations in the continuum [34, 35], or very explicitly on the lattice [34, 42]. The lattice formulation, which has the nice feature of being an explicitly cut-off theory, has the disadvantage of giving a simple theory in the dual variables only in the Villain approximation. The usual arguments (i.e. universality) employed to support the claim that the Villain approximation becomes exact in the continuum limit, do not go through so easily in the CP^1 case. Furthermore, we could see no practical advantage in using the lattice formulation here. Hence we shall discuss only the much simpler, though somewhat formal, continuum approach.

Let us denote by $J(x)$ the external charge density:

$$J(x) = \Sigma\, q_j \delta^2(x - x_j) \ . \tag{2.20}$$

We also introduce a θ-parameter in the usual way by the replacement

$$\frac{1}{g^2}\ S \longrightarrow \frac{1}{g^2}\ S + i\theta\hat{Q} \ . \tag{2.21}$$

The integration over ϕ is replaced by an integration over Θ_μ which is constrained locally by $\partial_\mu \Theta_\mu = 2\pi J$ (2.13). This change of variables involves only a linear transformation, and therefore the associated Jacobian has no field dependence and is just a number. The summation over all possible charge densities will be formally denoted by $\sum_J$. The constraint $\partial_\mu \Theta_\mu = 2\pi J$ is imposed by a Lagrange multiplier $B(\vec{x})$. The partition function becomes:

$$\begin{aligned} \mathcal{Z} = \sum_J \int \prod_x [dB]\ \prod_x [d\Theta_\mu]\ \prod_x [d\rho]\ \prod_x \left[\frac{df^2}{d\rho}\right] \times \\ \exp\left\{ - \frac{1}{g^2} \int [\, f^2(1-f^2)\ \Theta_\mu^2 + \left(\frac{df}{d\rho}\ \frac{1}{\sqrt{1-f^2}}\right)^2 (\partial_\mu\rho)^2 + \right. \\ \left. + iB(\partial_\mu\Theta_\mu - 2\pi J)] \ + \frac{i\theta}{\pi} \int d^2x \Theta_\mu f \frac{df}{d\rho}\, \partial_\mu\rho \right\}. \end{aligned} \tag{2.22}$$

Now the smooth magnetic background may be eliminated by integrating Θ_μ :

$$\mathcal{Z} = \sum_J \int \prod_x [dB] \prod_x [d\rho] \prod_x \left[\frac{df^2/d\rho}{f^2(1-f^2)}\right]' \times$$

$$\exp\left\{-\frac{1}{g^2}\int\left[\left(\frac{df}{d\rho}\right)^2 \frac{(\partial_\mu\rho)^2}{1-f^2} + \frac{1}{4}\,\frac{(\partial_\mu B - \frac{\theta g^2}{\pi} f \frac{df}{d\rho}\partial_\mu\rho)^2}{f^2(1-f^2)}\right]\right.$$
$$\left. - \frac{2\pi i}{g^2}\int BJ\right\}. \tag{2.23}$$

Eq. (2.23) contains a field theory of charges (vortices) immersed in a dielectric medium. The dielectric constant has a quite complicated distribution law.

If one wishes to get rid of the Jacobian in Eq. (2.23), the function f may be specified by demanding

$$\frac{df}{d\rho} = C\, f\, (1 - f^2)\ . \tag{2.24}$$

A possible solution (with C = 1) is given by:

$$f = \frac{1}{\sqrt{1+e^{-2\rho}}} \qquad -\infty < \rho < \infty\ , \tag{2.25}$$

which yields the following action,

$$S = -\frac{1}{g^2}\int\left[\frac{(\partial_\mu\rho)^2}{4\cosh^2\rho} + \cosh^2\rho\left(\partial_\mu B + \frac{\theta g^2}{2\pi}\,\partial_\mu\,\frac{1}{1+e^{2\rho}}\right)^2\right]$$
$$-\frac{2\pi i}{g^2}\int BJ\ . \tag{2.26}$$

For $\theta = 0$ we observe that the discrete symmetry of Eq. (1.1) (for N=2),

$z_1 \leftrightarrow z_2^*$, now appears as $\rho \leftrightarrow -\rho$. The range of integration of ρ is now infinite.

Assuming that the only relevant contributions have $q_j = \pm 1$ [(see Eq. (2.20)], one may replace $\oint_J$ by a summation on ($\pm$) charges and integration over positions. The summations are restricted by neutrality, but this should be a result of dynamics and so does not need to be imposed. One finds,

$$\mathcal{Z} = \int \prod_x [dB] \prod_x [d\rho] \times$$

$$\exp\left\{ -\frac{1}{g^2} \int \left[\frac{1}{4ch^2\rho} (\partial_\mu\rho)^2 + ch^2\rho \left(\partial_\mu B + \frac{\theta g^2}{2\pi} \partial_\mu \frac{1}{1+e^{2\rho}} \right)^2 \right] + \lambda \int \cos\left(\frac{2\pi}{g^2} B \right) \right\} . \tag{2.27}$$

The parameter λ is the fugacity of the gas; it has dimensions of $[\text{mass}^2]$ and must be present on dimensional grounds. Hence, there is a spontaneously generated mass in the theory.

But λ requires regularization, and this may be done in a variety of ways. As usual, we assume that different methods of regularization give the same final result. A quite similar problem occurs in the Abelian Higgs model, and we refer the interested reader to ref. [34] for a detailed discussion which can be translated in an obvious way to our model. Here we shall merely mention some of the alternatives. Perhaps the most straightforward approach is to put the theory on a lattice and perform the duality transformation in the Villain approximation, as already discussed. Then λ is given by $\frac{1}{a^2}$ with a as the lattice spacing. In a continuum theory, a reasonable guess would be $\lambda \sim \Lambda^2$, where Λ is the ultraviolet cut-off needed to define the theory in any case.

Although this is a good zeroth-order approximation, SU(2)-invariance [or in general SU(N)] requires a more complicated, field-dependent λ. This can be handled by the standard Faddeev-Popov trick [34], whereby "smeared out" field configurations are introduced instead of the singular configurations we have dealt with so far, and the smearing distance is related to the universal cut-off Λ. A third method of regularization, somewhere between the lattice and continuum approaches, is to define the theory on two-dimensional space considered as an array of cells. The fields take an average value within each cell. In this case λ is proportional to the size of a cell. Thus by any of these methods, we see that it is possible to define λ in a precise manner. The same problem will arise in the next section for the case of general N; our discussion here applies with obvious generalization.

III. TOPOLOGICAL EXCITATIONS IN CP^{N-1}

For $N > 2$ the explicit solution of the constraint $\sum_{\alpha=1}^{N} |z_\alpha|^2 = 1$ leads to quite complicated expressions [16]. Fortunately, topological singularities may be introduced without explicitly solving the constraint. We may, therefore, quite easily generalize from the $N = 2$ case. Introducing the parametrization

$$z_\alpha = e^{i\theta\alpha} \rho_\alpha \ ; \ \sum_{\alpha=1}^{N} \rho_\alpha^2 = 1 \ , \tag{3.1}$$

we rewrite the action (1.1) in the new variables:

$$S = \sum_\alpha \int (\partial_\mu \rho_\alpha)^2 + \sum_\alpha \int (\partial_\mu \theta_\alpha)^2 \rho_\alpha^2 - \sum_{\alpha\beta} \int \partial_\mu \theta_\alpha \partial_\mu \theta_\beta \rho_\alpha^2 \rho_\beta^2 \ . \tag{3.2}$$

The topological singularities are now introduced via the following constraints:

$$\varepsilon_{\mu\nu} \partial_\mu \partial_\nu \theta_\alpha = 2\pi \sum_i q_i^\alpha \, \delta^2(x - x_i^\alpha) \equiv J_\alpha(x), \ q_i^\alpha \in Z. \tag{3.3}$$

Since we did not, as yet, explicitly eliminate the gauge degree of freedom [as we did in the CP^1 case (2.7)] , it is clear that not all "charges", represented by the integers q_i^α , are physical. In order to get a feeling for the interaction between the singularities let us, for the moment, freeze the "dielectric medium" by looking at configurations with $\rho_\alpha = \frac{1}{\sqrt{N}}, \forall \alpha$. By doing this we have introduced some infinities resulting from the ultraviolet divergent self-energies, but we shall not worry about these. We are interested only in the mutual interactions, and singular self-energies could be taken care of by the local behavior of the medium. Thus we find the following effective action,

$$S_{eff} = \frac{1}{N} \int \partial_\mu \theta_\alpha X_{\alpha\beta} \partial_\mu \theta_\beta , \quad X_{\alpha\beta} = \delta\alpha\beta - \frac{1}{N} . \tag{3.4}$$

The numerical matrix $\underset{\sim}{X}$ is easily diagonalized: it has one zero eigenvalue and N-1 eigenvalues equal to one. Let the eigenvectors be denoted by $\vec{x}_{(\gamma)}$, $\gamma = 1, \ldots N$,

$$X \vec{x}_{(1)} = 0 \; ; \; X \vec{x}_{(\gamma)} = \vec{x}_{(\gamma)}, \quad \gamma > 1 ,$$

$$\vec{x}_{(\sigma)} \cdot \vec{x}_{(\tau)} = \delta_{\sigma\tau} . \tag{3.5}$$

It is easy to see that

$$x^{\alpha}_{(1)} = \frac{1}{\sqrt{N}} , \quad \alpha = 1, \ldots N , \tag{3.6}$$

and that $\underset{\sim}{X}$ is diagonalized by the matrix $\underset{\sim}{R}$ defined by

$$R = (\vec{x}_{(1)}, \ldots \vec{x}_{(N)}) , \; R^T R = 1, \; R^T XR = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} . \tag{3.7}$$

The appearance of the zero eigenvalue is a result of gauge invariance. Defining

$$\theta'_\alpha = R_{\beta\alpha} \theta_\beta \tag{3.8}$$

one gets

$$S = \frac{1}{N} \sum_{\gamma=2}^{N} \int \; (\partial_\mu \theta'_\gamma)^2 , \tag{3.9}$$

and $\theta_1' = \frac{1}{\sqrt{N}} \sum_\beta \theta_\beta$ has disappeared as it should. Topological singularities which appear only in θ_1' are physically irrelevant in exactly the same way as in the N = 2 case. Thus singular instantons have zero action here also.

To understand better the interaction of the original charges [only those we know how to quantize (3.3)], let us find the criterion for neutrality. A system of charges will be defined to be neutral if its total action is not infrared divergent. This means that for $\gamma \geq 2$ we must have usual neutrality in each $\partial_\mu \theta_\alpha'$(3.9). By (3.8), (3.7), and (3.3) this forces

$$\sum_\alpha x^\alpha_{(\gamma)} \sum_i \; q_i^\alpha = 0, \quad \gamma = 2, \ldots N$$

$$\Rightarrow \sum_i \; q_i^\alpha = \; kx^\alpha_{(1)} = k' \;, \forall \alpha \;. \tag{3.10}$$

Eq. (3.10) has of course many solutions, among which we find "N-pole" configurations where $k' = 1$ and there is exactly one singularity per component. Instantons and anti-instantons are of this latter type. The existence of other configurations with infrared-finite action suggests the possibility of having finite-action solutions to the Euclidean equations of motion other than instantons or anti-instantons. Examples of such solutions, obtained, for example, by embedding real O(3) instantons into CP^{N-1}, are indeed known to exist, but they are unstable [12, 43]. The pure instanton solutions are very simple when parametrized by the locations of the charges, a_α. Using notations similar to (2.2) we have for one instanton,

$$z_\alpha^{\text{inst}} \; (s) = c_\alpha \, (s - a_\alpha), \quad \alpha = 1, \ldots N \;. \tag{3.11}$$

Of course, one can also trivially generalize (2.1), but our whole point is that this is not very useful; it may even be misleading since it is strongly tied to the dilute gas approach.

In passing, we would like to observe the amusing fact that the quantization condition for the singularities in the θ'_γ-variables, q'_γ, may be expressed in the following way. Let $\lambda^{(\gamma)}$, $\gamma = 1,\ldots N-1$, be the subset of diagonal matrices of the standard representation of the SU(N)Lie algebra, with the slightly unusual normalization,

$$\mathrm{Tr}\ \lambda^{(A)}\ \lambda^{(B)} = \delta^{AB}\ . \tag{3.12}$$

Then the quantization condition reads

$$\left[\exp\left(2\pi i \sum_{\gamma=2}^{N} \lambda^{(\gamma-1)}\ q'_\gamma\right)\right]^N = \mathbb{1}\ . \tag{3.13}$$

In other words, the Lie algebra element $2\pi \sum_{\gamma=2}^{N} \lambda^{(\gamma-1)}\ q'_\gamma = Q'$ generates a matrix in the center of the group, Z(N). Two sets of singularities defined by Q'_1 and Q'_2 will interact (3.9) with a coefficient proportional to $\mathrm{tr}(Q'_1 Q'_2)$.

Exactly as in the case N = 2, the relevance of the positions of the topological singularities is sustained by the form of the long-range part of the instanton - anti-instanton interaction. To see this we shall use the method of refs. [20,22], and the reader is referred there for details. As a first step we rewrite the action in terms of unconstrained and gauge-invariant variables. We use the inhomogeneous coordinates for CP^{N-1} introduced in ref. [13]:

$$u_\alpha = \frac{z_\alpha}{z_N}, \ \alpha = 1, \dots N-1, \quad z_N \neq 0 \ . \tag{3.14}$$

The action of (1.1) is given by [13],

$$S = \sum_{\alpha,\beta} \int H_{\alpha\beta} \, \partial_\mu u^*_\alpha \, \partial_\mu u_\beta \, d^2x \equiv \int \mathcal{L} d^2x \ ,$$

$$H_{\alpha\beta} = \frac{\delta_{\alpha\beta} - \dfrac{u_\alpha u^*_\beta}{1 + \sum_\gamma |u_\gamma|^2}}{1 + \sum_\gamma |u_\gamma|^2} \ . \tag{3.15}$$

The instanton configuration is given by

$$u^{(1)}_\alpha = c_\alpha \frac{s - a^{(1)}_\alpha}{s - a^{(1)}_N} \quad ; \ a^{(1)}_N \ , a^{(1)}_\alpha \ \varepsilon \ \mathbb{C} \ , \alpha = 1, \dots, N-1, \tag{3.16}$$

and the anti-instanton by

$$u^{(2)}_\alpha = c_\alpha \left(\frac{s - a^{(2)}_\alpha}{s - a^{(2)}_N} \right)^* \quad ; \ a^{(2)}_N, \ a^{(2)}_\alpha \ \varepsilon \ \mathbb{C} \ ,$$

$$\alpha = 1, \dots, N-1 \ . \tag{3.17}$$

The constants c_α must be the same in order to allow smooth matching at infinity. For convenience we choose the instanton to be located at s = 0. With the assumption that there exists a distance R such that,

$$|a^{(1)}_\beta| << R << \frac{1}{N} \ |\sum_{\gamma=1}^{N} a^{(2)}_\gamma| \equiv |s_2|,$$

$$|a^{(2)}_\beta - a^{(2)}_{\beta'}| << |s_2| \ , \quad \beta,\beta' = 1 \ , \dots, N \ , \tag{3.18}$$

one can show [20, 22] that the interaction is given by

$$\delta S = \oint_{|x| = R} dx_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu u_\alpha)} \bigg|_{u_\alpha = u_\alpha^{(1)}} (u_\alpha^{(2)} - c_\alpha) + \text{c. c.} \tag{3.19}$$

Defining

$$b_\alpha^{(A)} = a_N^{(A)} - a_\alpha^{(A)} \, , \qquad \alpha = 1, \ldots, N-1, \qquad A = 1,2 \, , \tag{3.20}$$

we get

$$\delta S \approx \frac{2\pi}{s_2^2} \frac{1}{1 + \sum_{\gamma=1}^{N-1} |c_\gamma|^2} \times$$

$$\left[\sum_{\alpha=1}^{N-1} |c_\alpha|^2 \, b_\alpha^{(1)} \, b_\alpha^{(2)} - \frac{1}{1 + \sum_{\gamma=1}^{N-1} |c_\gamma|^2} \sum_{\alpha=1}^{N-1} |c_\alpha|^2 \, b_\alpha^{(1)} \times \right.$$

$$\left. \sum_{\alpha'=1}^{N-1} |c_{\alpha'}|^2 \, b_{\alpha'}^{(2)} \right] + \text{c. c.} \tag{3.21}$$

In order to compare with the naive prediction one gets from (3.4), we must make sure that the "vacuum" is given by $\rho_\alpha^2 = \frac{1}{N}$, $\alpha = 1$, ... N. This amounts to the demand $c_\gamma = 1$; $\gamma = 1$, ... N-1. Eq. (3.21) then gives

$$\delta S = \frac{2\pi}{N} \frac{1}{s_2^2} \left[\sum_{\alpha=1}^{N-1} b_\alpha^{(1)} \, b_\alpha^{(2)} - \frac{1}{N} \left(\sum_{\alpha=1}^{N-1} b_\alpha^{(1)} \right) \left(\sum_{\alpha=1}^{N-1} b_\alpha^{(2)} \right) \right] +$$

$$+ \text{c.c.} \tag{3.22}$$

In order to make contact with (3.9) we need the expression for the electrostatic dipole-dipole interaction between two neutral systems in two dimensions. Let one of the systems generate a potential (complex variables are used throughout)

$$\phi(s) = \frac{1}{2\pi} \sum_{\alpha=1}^{N} \bar{q}_\alpha \ln |s - a_\alpha| , \quad \sum_{\alpha=1}^{N} \bar{q}_\alpha = 0 . \tag{3.23}$$

The other system has charges q_α situated at b_α , $\alpha = 1,\ldots N$, $\Sigma q_\alpha = 0$. The interaction is

$$\delta S \sim \frac{1}{2\pi s^2} \quad P_a P_b \quad + \text{c.c.}, \tag{3.24}$$

where s is the distance between the systems, and the dipole moments are given by

$$P_a = \sum_{\alpha=1}^{N} \bar{q}_\alpha a_\alpha, \qquad P_b = \sum_{\alpha=1}^{N} q_\alpha b_\alpha .$$

From Eqs. (3.3, 3.4, 3.7-3.9) one can obtain expressions for $P_{a,b}$ appropriate to the system of vortices representing an instanton and anti-instanton. We thus find that the interaction is

$$\delta S = \frac{4\pi^2}{2\pi N s^2} \quad \sum_{\gamma=2}^{N} \sum_{\alpha=1}^{N} R_{\alpha\gamma} a_\alpha^{(1)} \sum_{\beta=1}^{N} R_{\beta\gamma} a_\beta^{(2)} + \text{c.c.} \quad . \tag{3.25}$$

Since

$$\sum_{\gamma=2}^{N} R_{\alpha\gamma} R_{\beta\gamma} = \delta_{\alpha\beta} - \frac{1}{N} \quad , \tag{3.26}$$

we get

$$\delta S = \frac{2\pi}{N s^2} \left[\sum_{\alpha=1}^{N} a_\alpha^{(1)} \ a_\alpha^{(2)} - \frac{1}{N} \left(\sum_{\alpha=1}^{N} a_\alpha^{(1)} \right) \left(\sum_{\beta=1}^{N} a_\beta^{(2)} \right) \right] + \text{c.c.} \ . \tag{3.27}$$

It is now trivial to show that this result completely agrees with (3.22).

Therefore the classical interaction between an instanton and an anti-instanton may be viewed as resulting from the vorticity structure. Since the interaction between two instantons is zero we know that, exactly as in CP^1, the medium may not always be neglected. Nevertheless, since instanton-anti-instanton configurations have more entropy it may ultimately turn out that the medium is not very important for the most probable configurations.

We now go on to derive a field theory for the vortices introduced by the constraints in (3.3). With these constraints incorporated the partition function of the system is given by:

$$\mathcal{Z} = \int [\pi\rho_\alpha d\rho_\alpha]\,[\delta(\Sigma\rho_\alpha^2 - 1)][\pi dC_\mu^\alpha] \; \sum_{J_\alpha} \int [\pi dB_\alpha]$$
$$e^{-\frac{1}{g^2}\left[\sum_\alpha \int (\partial_\mu\rho_\alpha)^2 + \sum_\alpha \int \left(C_\mu^\alpha\right)^2 \rho_\alpha^2 - \sum_{\alpha,\beta} \int C_\mu^\alpha C_\mu^\beta \rho_\alpha^2 \rho_\beta^2\right]}$$
$$e^{-\frac{i}{g^2}\sum_\alpha \int B_\alpha \left(\partial_\mu C_\mu^\alpha - J_\alpha\right)} \quad , \tag{3.28}$$

where we introduced the Lagrange multipliers B_α and denoted $\partial_\mu \theta_\alpha$ by $\varepsilon_{\mu\nu} C_\mu^\alpha$. The diagonalization of the quadratic form in C_μ^α is made easier by rescaling

$$C_\mu^\alpha \longrightarrow C_\mu^{\alpha'} = \rho_\alpha C_\mu^\alpha \quad . \tag{3.29}$$

The effective action is now

$$S = \sum_{\alpha} \int (\partial_\mu \rho_\alpha)^2 + \sum_{\alpha,\beta} \int C_\mu^{\prime\alpha} X_{\alpha\beta}(\rho) C_\mu^{\prime\beta} +$$

$$+ i \sum_{\alpha} \int B_\alpha J_\alpha + i \sum_{\beta} \int \frac{1}{\rho_\alpha} \partial_\mu B_\alpha C_\mu^{\prime\alpha} , \tag{3.30}$$

$$X_{\alpha\beta}(\rho) = \delta_{\alpha\beta} - \rho_\alpha \rho_\beta \equiv (\mathbb{1} - P)_{\alpha\beta} . \tag{3.31}$$

The constraint on ρ_α makes P a projection operator properly normalized (i.e. $P^2 = P$). The matrix $\underset{\sim}{X}$ again has exactly one zero eigenvalue, all the other eigenvalues being equal to one. The zero eigenvector $x_{o\alpha}$ is given by

$$x_{o\alpha} = \rho_\alpha , \qquad \alpha = 1, \ldots N . \tag{3.32}$$

Let $\underset{\sim}{R}$ be the orthogonal matrix which diagonalizes $\underset{\sim}{X}$:

$$R^T R = \mathbb{1} , \qquad R^T X R = \begin{pmatrix} 0 & & 0 \\ & 1 & \\ 0 & & \ddots \; 1 \end{pmatrix} , \tag{3.33}$$

$$R_{\alpha 1} = x_{o\alpha} , \quad \alpha = 1, \ldots N .$$

Transforming variables again to

$$C_\mu^{\prime\prime\alpha} = \sum_{\beta=1}^{N} R^{\beta\alpha} C_\mu^{\prime\beta} , \qquad \alpha = 1, \ldots N , \tag{3.34}$$

and performing the integration over $C_\mu^{\prime\prime\alpha}$ (which eliminates the "smooth magnetic background") we arrive at

$$\mathcal{Z} = \sum_{J_\alpha} \int \left[\pi \frac{d\rho_\alpha}{\rho_\alpha}\right] \left[\delta\left(\sum_\alpha \rho_\alpha^2 - 1\right)\right] \left[\pi dB_\alpha\right] \left[\delta\left(\sum_\alpha B_\alpha\right)\right]$$

$$\exp\left\{ -\frac{1}{g^2}\left[\sum_\alpha \int (\partial_\mu \rho_\alpha)^2 + \frac{1}{4}\sum_\alpha \int \frac{1}{\rho_\alpha^2} (\partial_\mu B_\alpha)^2 - i \sum_\alpha \int B_\alpha J_\alpha\right]\right\} . \tag{3.35}$$

If we restrict $\sum_{J_\alpha}$ in a convenient manner, we arrive at the analogue of (2.27):

$$\mathcal{Z} = \int \left[\pi \frac{d\rho_\alpha}{\rho_\alpha}\right] [\delta(\sum_\alpha \rho_\alpha^2 - 1)] [\pi dB_\alpha] [\delta(\sum_\alpha B_\alpha)] \times$$

$$\exp\left\{ -\frac{1}{g^2}\left[\sum_\alpha \int (\partial_\mu \rho_\alpha)^2 + \frac{1}{4}\sum_\alpha \int \frac{1}{\rho_\alpha^2} (\partial_\mu B_\alpha)^2\right] + \lambda \sum_\alpha \int \cos \frac{2\pi}{g^2} B_\alpha \right\} . \tag{3.36}$$

One can also introduce a θ-parameter in the same manner as was done for CP^1. The action is changed by the addition of δS_θ :

$$\delta S_\theta = \frac{i\theta}{2\pi} \int q(x) d^2x = \frac{i\theta}{\pi} \sum_{\alpha,\beta} \int d^2x \partial_\mu \rho_\alpha R_{\alpha\beta} C''^{\beta}_\mu , \tag{3.37}$$

where $q(x)$ is the topological charge density, and $\underset{\sim}{R}$ and $\vec{C}''_\mu$ have been defined in (3.33) and (3.34). The extra term δS_θ generates the following change in (3.35):

$$\partial_\mu B_\alpha \longrightarrow \partial_\mu \left(B_\alpha + \frac{g^2\theta}{2\pi} \rho_\alpha^2\right) . \tag{3.38}$$

Eq. (3.36) now becomes (after a shift in B_α),

$$\mathcal{Z}_{\theta} = \int \left[\pi \frac{d\rho_{\alpha}}{\rho_{\alpha}}\right] \left[\delta\left(\sum_{\alpha} \rho_{\alpha}^{2} - 1\right)\right] \left[\pi dB_{\alpha}\right]$$

$$\left[\delta\left(\sum_{\alpha} B_{\alpha} - \frac{1}{2\pi} g^{2}\theta\right) \exp\left\{-\frac{1}{g^{2}} \left[\sum_{\alpha} \int (\partial_{\mu}\rho_{\alpha})^{2} + \right.\right.\right.$$

$$\left.\left. + \tfrac{1}{4} \sum_{\alpha} \int \frac{1}{\rho_{\alpha}^{2}} (\partial_{\mu}B_{\alpha})^{2}\right] + \lambda \sum_{\alpha} \int \cos\left(\frac{2\pi}{g^{2}} B_{\alpha} - \theta\rho_{\alpha}^{2}\right)\right\}. \quad (3.39)$$

IV. TOPOLOGICAL EXCITATIONS AND THE 1/N EXPANSION.

There are two arguments that support the belief that topological excitations are unimportant in the CP^{N-1} models. The first argument runs as follows. Cluster decomposition is violated by a pure instanton (or anti-instanton) gas, and therefore one must have a mixed gas with fluctuating "dielectric medium". So far, one has been able to mix up instantons and anti-instantons only in the dilute gas approximation. However, Fateev, Frolov, and Schwarz [13, 23], and Berg and Lüscher [12] have performed the lowest order quantum expansion around exact solutions to the classical equations and have shown that the instanton quarks (merons) are liberated from their bound state, the instanton. As a result, a quark (meron) plasma is formed, which is very different from the dilute gas mentioned above. As we shall see, this section also supports the idea of a quark plasma. It is necessary, however, to go beyond an expansion around exact classical solutions in order to avoid violating cluster decomposition. We shall investigate this question in the next section and present several models of an acceptable quark plasma.

In this section, we will confront the second argument, which relies on the large N limit. If the large N limit is a good approximation for all $N \geqslant 2$, as it seems to be, and if the contribution of an instanton is suppressed by a barrier penetration factor of the form

$$\exp\left(-\frac{\text{const.}}{g^2}\right) = \exp\left(-\text{const.}\ N\right) , \tag{4.1}$$

then instantons disappear as $N \to \infty$ and they cannot be important for finite N [4,32,33] . The exponent that appears in the barrier penetration factor is the classical action of the instanton, which goes like $\frac{1}{g^2} \approx N$. This argument may be valid for a dilute instanton gas, in which the N quarks that make up one instanton are tightly bound. If, however, these quarks are liberated to form a plasma, it is then plausible that the expression analogous to (4.1) per each quark will have an exponent reduced by a factor of N, so their contribution will not disappear in the large N limit. Indeed, in what follows, we shall argue in favor of a quark plasma as opposed to a dilute instanton gas. We shall also show that in the plasma phase the model confines charge and that the confinement is entirely due to the plasma of topological excitations (quarks), which do not disappear in the large N limit.

We start with the criterion for confinement [11],

$$\mathcal{Z}_\theta^{-1} \frac{d^2\mathcal{Z}_\theta}{d\theta^2} \neq 0 . \tag{4.2}$$

This criterion follows by noticing that there must be a non-vanishing derivative of the expectation value of the topological density in order for the photon propagator (in general the propagator of the topological gauge field) to have a pole at $p^2 = 0$. This criterion is essentially equivalent to that for the Wilson loop [11]. The expression for $\mathcal{Z}_\theta$ is given by Eq. (3.39), but prior to taking the large N-limit it is necessary to scale several of the variables there as follows:

$$g^2N = \kappa , \quad \theta = \bar{\theta} N, \; \rho_\alpha = g\,\bar{\rho}_\alpha , \quad B_\alpha = g^2\left(\bar{B} + \frac{\bar{\theta}}{2\pi}\right). \tag{4.3}$$

It is also convenient to exponentiate the delta function constraints of (3.39) by means of two Lagrange multipliers u and v. As a result, we have

$$\mathscr{Z}_\theta = \int [du]\,[dv] \left[\Pi \frac{d\bar{\rho}_\alpha}{\bar{\rho}_\alpha}\right] [\Pi d\bar{B}]_\alpha \;\exp\left\{\int d^2x \sum_\alpha \left[-(\partial_\mu \bar{\rho}_\alpha)^2 \right.\right.$$

$$- \tfrac{1}{4} \frac{1}{\bar{\rho}_\alpha^{\,2}} (\partial_\mu \bar{B}_\alpha)^2 + \lambda \cos\,[2\pi\bar{B}_\alpha - \bar{\theta}\,(\kappa\,\bar{\rho}_\alpha^{\,2} - 1)]$$

$$\left.\left. + iu\,\bar{\rho}_\alpha^{\,2} + iv\bar{B}_\alpha\right] - i\,\frac{N}{\kappa} \int d^2x\; u\right\}. \qquad (4.4)$$

This expression can be rewritten as

$$\mathscr{Z}_\theta = \int [du]\,[dv] \quad \exp\left\{ N\;[\log I(u,v,\bar{\theta}) - \frac{i}{\kappa}\int d^2x\; u]\right\} , \qquad (4.5a)$$

where,

$$I = \int \left[\frac{d\bar{\rho}}{\bar{\rho}}\right] [d\bar{B}] \quad \exp\left\{ \int d^2x \;\left[-(\partial_\mu\,\bar{\rho})^2 - \tfrac{1}{4}\frac{1}{\bar{\rho}^{\,2}} (\partial_\mu\,\bar{B})^2 \right.\right.$$

$$\left.\left. + \lambda \cos\,[2\pi\bar{B} - \bar{\theta}(\kappa\,\bar{\rho}^{\,2} - 1)] + iu\bar{\rho}^{\,2} + iv\bar{B}\right]\right\} . \qquad (4.5b)$$

The large N limit is now given by the standard saddle point method, that is, the integrand in (4.5) becomes as $N \to \infty$

$$\exp\left\{ N \;\; [\log I\;(u_o,\, v_o,\, \bar{\theta}) - \frac{i}{\kappa}\int d^2x\; u_o]\right\} , \qquad (4.6a)$$

where the saddle points u_o and v_o are determined by the equations

$$\frac{\partial I}{\partial u_o} = \frac{i}{\kappa}\, I, \qquad \frac{\partial I}{\partial v_o} = 0. \qquad (4.6b)$$

Eqs. (4.5) and (4.6a) can be combined into the following expression for the left-hand side of (4.2):

$$\mathcal{Z}_\theta^{-1} \frac{d^2\mathcal{Z}_\theta}{d\theta^2} = \frac{V}{N} \frac{d^2}{d\bar{\theta}^2} \left(\frac{1}{V} \log I - \frac{i}{\kappa} u_o \right), \tag{4.7}$$

where V = volume (area) of space, and $\frac{d}{d\bar{\theta}} = \frac{\partial}{\partial\bar{\theta}} + \frac{\partial u_o}{\partial\bar{\theta}} \frac{\partial}{\partial u_o} + \frac{\partial v_o}{\partial\bar{\theta}} \frac{\partial}{\partial v_o}$, since the saddle points implicitly depend on $\bar{\theta}$. A disconnected contribution to (4.7) vanishes since $\mathcal{Z}_\theta$ is an even function of θ [see (4.4)] . For the sake of definiteness, we exhibit the first term in (4.7),

$$\frac{1}{V} \left(\frac{\partial^2}{\partial\bar{\theta}^2} \log I \right)_{\bar{\theta}=0} = - \frac{\lambda}{IV} \int d\bar{\rho} d\bar{B} \left[\int d^2x (\kappa\bar{\rho}^2 - 1)^2 \cos(2\pi\bar{B}) \right]$$

$$\exp \left\{ \int d^2x \left[- (\partial_\mu\bar{\rho})^2 - \tfrac{1}{4} \frac{(\partial_\mu\bar{B})^2}{\bar{\rho}^2} + \lambda \cos(2\pi\bar{B}) + iu_o\bar{\rho}^2 \right.\right.$$

$$\left.\left. + iv_o\bar{B} \right] \right\} \equiv -\lambda \langle (\kappa\bar{\rho}^2 - 1)^2 \cos[2\pi\bar{B}(0)] \rangle . \tag{4.8}$$

So we conclude that in the large N limit the topological excitations do not go away; the term $\lambda \cos(2\pi\bar{B})$ which sums up these excitations survives in (4.5b). Indeed, it is easy to see that the topological charge density (and so the Wilson loop) is proportional to λ [see (4.8)]. Therefore, if topological excitations are projected out of the functional integral by setting $\lambda = 0$, the Wilson loop would vanish and there would be no θ-dependence. It is natural to suppose that λ determines the mass scale of the spectrum of the model. In that case, the N-dependence given by (4.7) is in agreement with the result of the direct $\frac{1}{N}$ calculation [10,11] .

It is still not clear that there is a non-vanishing θ-dependence since matrix elements of the type given by (4.8) could vanish. Barring an unlikely accidental cancellation, in two dimensions this can happen because of a long-range correlation and the resulting infrared singularity in correlation functions. In fact, as has been shown by Coleman [39], in two dimensions a matrix element of the form

$$\langle \exp [2\pi i c_1 \phi(0)] \rangle \tag{4.9a}$$

vanishes if the correlation function of ϕ is that of a free massless field:

$$\langle \phi(0)\phi(x) \rangle = c_2 \log (x^2) . \tag{4.9b}$$

Such a long-range contribution in the spectrum of the field $\bar{\rho}^2$ would imply the existence of a massless particle. Since it is known that a mass term proportional to ρ^2 is generated, we instead focus our attention on the field $\bar{B}$. One then has to distinguish between two possible phases of the gas formed out of topological excitations (quarks). In the first phase, N quarks are bound together in a neutral system and these clusters make up a dilute gas. For N = 2, this is the well-known [36] dipole phase, where oppositely charged quarks pair up. There is clearly long-range correlation in this phase due to the Coulomb force, and the matrix element of Eq. (4.8) vanishes since

$$\langle \cos (2\pi\bar{B})\rangle = 0. \tag{4.10}$$

A similar argument applies to the other terms in Eq. (4.7), and we get no θ-dependence and no charge confinement in this phase.

On the other hand, if the quark gas is in a plasma phase, the long-range force is screened and there is no singularity of the form (4.9b) in the matrix elements; the expectation value of $\cos (2\pi\bar{B})$ does not vanish and consequently there is confinement. Since the topological charge density has already been calculated [10] and found to be non-zero in the large N limit, we can reverse our argument and consider it as very strong evidence in favor of the plasma phase for the quarks.

To sum up, we conclude that topological excitations survive in the large N limit, they are entirely responsible for confinement, and they form a gas in its plasma phase. Finally, we stress that our purpose in this section was not to do explicit large N calculations, which can be done much more easily without first carrying out the duality transformation. Rather, we exploited the duality transformation and the well-known large N results to learn about topological excitations in the large N limit.

V. TWO COMPONENT PLASMA MODEL

In the previous sections we have argued that topological excitations or vortices are important and must be taken into account in the functional integral of the CP^{N-1} models. An advantage of the field theories derived from the dual transformations is that they allowed us to assess the importance of the vortices without knowing too much about the dielectric medium. However, we do know that the statistically important configurations, which satisfy cluster decomposition, are those with a medium which fluctuates between the charges. We may describe this by allocating a spin variable to each vortex (meron), and then having the spins distributed essentially randomly. But now we have to know something about the spin-spin interaction in order to study the statistical mechanics of this plasma of charges (vortices) which carry a spin label. In particular, we are interested in the possible phases of this plasma. Since, so far, we have been unable to isolate the spin-spin interaction directly from the field theory, we present in this section a simplified model (somewhat along the lines of ref.[22]) of the vortex plasma, which will provide some guidance about the phases. For simplicity we discuss only the CP^1 model.

We wish to consider a two-dimensional gas of particles of two types, q and p. Each type of particle carries a positive or negative charge which interacts via a Coulomb potential. In addition, there is a repulsive short-range Yukawa interaction between p and q-type particles. Hence, the model describes the gas of instanton quarks or vortices — type q particles, and anti-instanton quarks — type p particles, with their known Coulomb interactions. We should note that these Coulomb interactions have different sources. The Coulomb interaction between instanton (or anti-instanton) quarks by themselves is a quantum effect [12, 23] ; the instanton – anti-instanton Coulomb (dipolar) interaction (and the absence of an instanton-instanton interaction) is a classical result [20,22]. In our model we shall take this distinction into account only by assigning a different strength for the p-q interaction than for the q-q (p-p) interaction. In addition to the different strength for the p-q Coulomb interaction (which partially takes into account the spin-spin interactions), we have also approximated the spin-spin interactions by the Yukawa repulsion (which is technically convenient) only between an instanton quark and an anti-instanton quark. This can be justified on the grounds that it overcomes the double-counting problem [44].

So we are led to consider the following grand partition function

$$\mathcal{Z} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{\lambda_p^{2m} \lambda_q^{2n}}{(m!)^2 (n!)^2} \int_V \prod^n da_i \int_V \prod^n db_i \int_V \prod^m dc_k \int_V \prod^m dd_k$$

$$\exp\left\{ \beta\, q^2 \left[\sum_{i\neq j} (\ln |a_i - a_j| + \ln |b_i - b_j|) \right.\right.$$

$$\left. - 2 \sum_{i,j} \ln |a_i - b_j| \right] +$$

$$+ \beta p^2 \left[\sum_{k\neq l} (\ln |c_k - c_l| + \ln |d_k - d_l|) - \right.$$

$$\left. - 2 \sum_{k,l} \ln |c_k - d_l| \right] +$$

$$\beta \varepsilon p q \sum_{i,k} (\ln a_i - c_k + \ln b_i - d_k \quad -$$

$$- \ln |a_i - d_k| - \ln |b_i - c_k|) \Big\} \quad \times$$

$$\exp \Big\{ - \beta\gamma \sum_{i,k} \big[K_o(m|a_i - c_k|) + K_o(m|a_i - d_k|) +$$

$$+ K_o(m|b_i - c_k|) + K_o(m|b_i - d_k|) \big] \Big\} . \tag{5.1}$$

In (5.1) the temperature is $\frac{1}{\beta}$, and we expect that in general it will be different from that found for the pure instanton-quark plasma [12,23] since there are now both kinds of quarks present with interactions between them. The fugacity, λ, is related to the chemical potential μ by $\lambda \propto e^{\beta\mu}$, and it has the dimensions of V^{-1}, where V is the volume (area) of integration in (5.1). The charge of the instanton quarks is q, and for anti-instanton quarks it is p. Eventually we shall set $p = q$ and $\lambda_p = \lambda_q$ in agreement with the discrete symmetry $w \to \frac{1}{w^*}$ already mentioned in Sec. II, but for bookkeeping purposes it is convenient to keep them different for a while. We have imposed neutrality within each component of the Coulomb gas as usual in two dimensions [45]. The strength of the Coulomb interaction between p- and q-type particles is measured by εpq. The Yukawa potential in two dimensions is given by $K_o(m|r_i - r_j|)$, which is the modified Bessel function of zeroth order, and γ in (5.1) measures its strength, while m determines the range.

In order to study this system we shall use the well-known Gaussian representation [22, 44-47] to express it as a field theory. To this end we rewrite (5.1) as follows:

$$\mathcal{Z} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{\lambda_p^{2m} \lambda_q^{2n}}{(m!)^2 (n!)^2} \underset{\sum_i^{2n} q_i=0}{\sum_{q_1=\pm q} \cdots \sum_{q_n=\pm q}} \underset{\sum_k^{2m} p_k=0}{\sum_{p_1=\pm p} \cdots \sum_{p_m=\pm p}}$$

$$\int_V d^2x_1 \ldots d^2x_{2n} \int_V d^2y_1 \ldots d^2y_{2m}$$

$$\exp \left\{ \beta \left[\sum_{i,j}^{2n} q_i q_j \ln|x_i - x_j| - \sum_i^{2n} q_i^2 \ln(0) + \right.\right.$$

$$+ \sum_{k,l}^{2m} p_k p_l \ln|y_k - y_l| - \sum_k^{2m} p_k^2 \ln(0) +$$

$$+ \varepsilon \sum_{i,k}^{n,m} q_i p_k \ln|x_i - y_k| -$$

$$\left.\left. - \gamma \sum_i^{2n} \sum_k^{2m} K_o(m\,|x_i - y_k|) \right] \right\} . \qquad (5.2)$$

In (5.2) we have explicitly subtracted the infinite Coulomb self-energy terms. Without subtracting these self-energy terms, the exponential in (5.2) can be expressed as a field theoretic functional integral,

$$\frac{1}{\mathcal{N}} \int \mathcal{D}\chi_+ \mathcal{D}\chi_- \mathcal{D}\phi_+ \mathcal{D}\phi_- \exp\left\{ -\tfrac{1}{2} \left[\int (\nabla\chi_+)^2 \, d^2x + \right.\right.$$

$$\left. + \int (\nabla\chi_-)^2 \, d^2x + \int (\nabla\phi_+)^2 \, d^2x + \int (\nabla\phi_-)^2 \, d^2x \right] -$$

$$- \frac{m^2}{2} \left[\int \phi_+^2 \, d^2x + \int \phi_-^2 \, d^2x \right] + \sqrt{\beta} \left[i \sqrt{1 + \frac{\varepsilon}{2}} \int \chi_+ \rho_+^c \, d^2x + \right.$$

$$+ i \sqrt{1 - \frac{\varepsilon}{2}} \int \chi_- \rho_-^c \, d^2x + i \sqrt{2\gamma} \int \phi_+ \rho_+ d^2x +$$

$$\left.\left. + \sqrt{2\gamma} \int \phi_- \rho_- \, d^2x \right] \right\} , \qquad (5.3)$$

with the densities

$$\rho_\pm^c = \sqrt{2\pi}\left[\sum_i^n q_i \delta^2(\vec{x} - \vec{x}_i) \pm \sum_k^m p_k \delta^2 (\vec{x} - \vec{y}_k)\right]$$

$$\rho_\pm = \sqrt{2\pi}\left[\sum_i^{2n} \delta^2 (\vec{x} - \vec{x}_i) \pm \sum_k^{2m} \delta^2 (\vec{x} - \vec{y}_k)\right], \tag{5.4}$$

and $\mathcal{N}$ is the functional integral with only terms quadratic in the fields in the exponent.

Since non-neutral plasmas do not contribute to the functional integral, because they are suppressed by infrared divergence, we can relax the neutrality conditions and then the result of doing the summations in (5.2) is

$$\mathcal{Z} = \frac{1}{\mathcal{N}} \int \mathcal{D}\chi_+ \; \mathcal{D}\chi_- \; \mathcal{D}\phi_+ \, \mathcal{D}\phi_-$$

$$\exp\left\{ - \tfrac{1}{2} \left[\int (\nabla\chi_+)^2 + \int (\nabla\chi_-)^2 + \int (\nabla\phi_+)^2 + m^2\phi_+^2\right.\right.$$

$$\left. + \int (\nabla\phi_+)^2 + m^2\phi_-^2\right] + 8\lambda \int d^2x \cos\left[\sqrt{2\pi\beta}\sqrt{1+\frac{\varepsilon}{2}}\, q\chi_+ (x)\right]$$

$$\cos\left[\sqrt{2\pi\beta} \sqrt{1 - \frac{\varepsilon}{2}}\; q\, \chi_- (x)\right] \exp\left[i \sqrt{2\pi\beta} \sqrt{2\gamma}\, \phi_+(x)\right]$$

$$\left.\cosh\left[\sqrt{2\pi\beta} \sqrt{2\gamma}\; \phi_- (x)\right] \right\} , \tag{5.5}$$

where we have set $p = q$, and $\lambda_q = \lambda_p = \lambda$. So the grand partition function of our model system has been expressed as a functional integral of four scalar fields. With this representation one can more easily study the possible ground states of the system, and so discover in which phases the two-component plasma can be. We should remark that not all the fields in (5.5) are to be integrated along the real axis, otherwise the integral diverges (and, as we shall see, all the fields being real is unphysical anyway). One can check that a contour exists where the functional integral makes sense at least formally.

The ground states of the field theory described by (5.5) can be found by searching for extrema of the potential

$$\mathcal{V} = \frac{m^2}{2} \left(\phi_+^2 + \phi_-^2\right) - 8\lambda \cos(a\chi_+) \cos(b\,\chi_-)$$

$$\exp(ic\phi_+) \cosh(c\phi_-),$$

with

$$a = \sqrt{2\pi\beta}\,\sqrt{1+\frac{\varepsilon}{2}}\,q,$$

$$b = \sqrt{2\pi\beta}\,\sqrt{1-\frac{\varepsilon}{2}}\,q,$$

$$c = \sqrt{2\pi\beta}\,\sqrt{2\gamma}\quad . \tag{5.6}$$

The variations with respect to χ_+ and χ_- give $\chi_\pm = 0,\ n\pi$ as the extrema. Using this, the result of variations with respect to ϕ_+ and ϕ_- may be written as the following coupled equations for the extrema of $\mathcal{V}$:

$$\phi_+ = \pm\, i8\lambda \frac{c}{m^2}\, e^{ic\phi_+} \cosh(c\phi_-), \tag{5.7a}$$

$$\phi_- = \pm\, 8\lambda \frac{c}{m^2}\, e^{ic\phi_+} \sinh(c\phi_-). \tag{5.7b}$$

Define

$$\Phi_q = c(i\phi_+ + \phi_-) \tag{5.8a}$$

$$\Phi_p = c(i\phi_+ - \phi_-) \quad . \tag{5.8b}$$

Then (5.7) can be rewritten as

$$\Phi_q = \pm\, k\, e^{\Phi p} \tag{5.9a}$$

$$\Phi_p = \pm\, k\, e^{\Phi q} \quad , \tag{5.9b}$$

with

$$k = 8\lambda \frac{c^2}{m^2} > 0 \quad . \tag{5.9c}$$

We shall not analyze in complete generality the coupled system of transcendental equations (5.9), nevertheless our study will be sufficient to identify what we believe to be physically relevant solutions.

To find out what kinds of solutions we should consider, we look at the average number of particles, $\overline{N}_q$ and $\overline{N}_p$. $\overline{N}$ is as usual calculated by

$$\overline{N} = \lambda \frac{\partial}{\partial\lambda} \ln \mathcal{Z}. \tag{5.10}$$

In order to determine $\overline{N}_q$ and $\overline{N}_p$ individually, we need $\mathcal{Z}$ in terms of λ_q and λ_p . The λ and Φ-dependent part of $\mathcal{Z}$ for $\lambda_q \neq \lambda_p$ is

$$\exp\left(\lambda_q \int e^{\Phi_q} + \lambda_p \int e^{\Phi_p}\right) . \tag{5.11}$$

Hence

$$\overline{N}_q = \lambda \langle \int e^{\Phi_q} \rangle \tag{5.12a}$$

$$\overline{N}_p = \lambda \langle \int e^{\Phi_p} \rangle \tag{5.12b}$$

where we have again set $\lambda_q = \lambda_p = \lambda$. If $\Phi_{\substack{p\\q}}$ is a purely imaginary solution to (5.9), we find $\overline{N}$ is not real, which is hardly respectable. The same type of objection would appear to apply to complex solutions. Hence we consider only purely real solutions to (5.9). The analysis is given in the appendix; the results are as follows:

$\Phi_q = \Phi_p = s$, $s = \pm k e^s$	$-$	$0<k<\infty$, 1 solution = $s<0$
	$+$	$0<k<e^{-1}$, 2 solutions: s_1, $s_2 > 0$
	$+$	$k = e^{-1}$, 1 solution: $s>0$
	$+$	$k>e^{-1}$, no solution

$\Phi_q \neq \Phi_p$	$-$	$0<k\leqslant e$, no solution
	$-$	$k>e$, 2 solutions: $d_1 \neq d_2$ a) $\Phi_q = d_1$, $\Phi_p = d_2$ b) $\Phi_q = d_2$, $\Phi_p = d_1$ $-d = k\, e^{-ke^{d}}$, $d<0$
	$+$	$0<k<\infty$, no solution

(5.13)

So we have found two general kinds of real solutions, $\Phi_q = \Phi_p$ and $\Phi_q \neq \Phi_p$, which are classical, constant fields at which the potential $\mathcal{V}$ (5.6) of our system of instanton and anti-instanton quarks is at an extremum, which in general will be a saddle point.

One can usually deform the contour of integration of the fields in such a way that one moves through the saddle in the optimum direction. Hence these solutions represent true ground states about which one can calculate quantum fluctuations.

To see what these solutions imply about the physics of our system, we again look at the average number of particles, $\overline{N}_q$ and $\overline{N}_p$ (5.12) . Not surprisingly, for $\Phi_q = \Phi_p = s$,

$$\overline{N}_q = \overline{N}_p = \lambda V e^s \quad . \tag{5.14}$$

So this solution is the symmetric ground state with on the average equal numbers of both kinds of particles.

For $\Phi_q \neq \Phi_p$, we have

$$\overline{N}_q = \lambda V e^{d_1} \text{ and } \overline{N}_p = \lambda V e^{d_2} , \tag{5.15}$$

or $d_1 \leftrightarrow d_2$. Thus this is the asymmetric ground state. In the limit that say $d_1 = 0$ which implies $d_2 = -\infty$ and $k = \infty$, there are only instanton quarks in this ground state. This is an extreme condition classically, namely $k = \infty$ implies $\gamma = \infty$ or $m = 0$, which we do not believe is realistic.

Thus two pictures of the instanton—anti-instanton quark plasma emerge. In one type of plasma phase, there is a symmetric mixture of both kinds of particles throughout space. The other type of phase is a droplet picture in which in some regions there are more instanton quarks, while in other regions there are more anti-instanton quarks. In this type of plasma phase, there may also be regions in which a symmetric mixture exists, since we found all three solutions exist when the asymmetric ones do. However, the asymmetric and symmetric solutions in general will not have the same energy. So one or the other will be the true ground state. The droplet phase can only exist for

$k = 32\,\lambda\pi \dfrac{\beta\gamma}{m^2} > e$. Unfortunately, we do not have independent information about these parameters, so the question of which of these pictures is correct cannot be decided.

It should be noted that there is also the well-known Coulomb dipole-plasma transition in our system, but from arguments given in section IV, we believe that the system is in the plasma phase.

VI. SUMMARY AND DISCUSSION

We have identified a class of Euclidean configurations which appear to be dominant in the functional integral of the CP^{N-1} models. Unlike the very restrictive set of exact classical solutions, they have sufficient entropy to be statistically important, and they also satisfy cluster decomposition. This class of topological excitations, which we have defined in a point-like manner, are closely related to the merons of the O(3) model, and they may be viewed as constituents of instantons. However, they are defined independently of instantons and exist even in theories which do not possess stable classical instantons (e.g. the O(N) non-linear sigma models). We have also discussed for CP^1 two different possibilities for a concrete statistical mechanical plasma of excitations. This treatment can be generalized to arbitrary N.

We have been able to check whether these topological excitations are truly important since the CP^{N-1} models have been solved in the $\frac{1}{N}$ expansion. We have shown not only that these configurations survive as $N \to \infty$, but that in the plasma phase they are responsible for the effects encountered within the $\frac{1}{N}$ expansion — confinement, θ-dependence, and dynamical mass generation. Since these objects carry fractional topological charge, our work here answers affirmatively the question whether field configurations with fractional topological charge contribute to the functional integral.

The CP^{N-1} models are interesting because they seem to provide a laboratory for techniques applicable, at least in principle, to QCD. Given the success of the quasi-classical approach here, we are encouraged that similar configurations may be important for the long-distance features of QCD. Of course, in CP^{N-1} we had the happy situation that the $\frac{1}{N}$ expansion could be effectively employed. Since this does not yet seem to be true for QCD, it is all the more important that an alternative method, the quasi-classical, can be tractable and effective. Nevertheless, if we take seriously the situation of CP^{N-1} as a prototype for that of QCD, then we will want to satisfy the criterion that field configurations, which we hope to be dominant for QCD, survive in the large N (color now) limit.

The idea of viewing QCD instantons as built out of elementary point-like constituents has been presented in ref. [48]. There this idea was used for the treatment of quantum fluctuations around exact multi-instanton configurations. Instead, we hope to generalize the approach we have taken in CP^{N-1}, where point-like topological excitations are not directly tied to exact classical solutions. Of course, it may also turn out that the system of interacting point-like excitations, even though it can survive as $N \to \infty$, is not the direct agent for confinement. It may assume only the role of providing the necessary background for other, no longer point-like, excitations, like the Z(N) fluxon surfaces [49]. In any case we feel that the extension of our work to QCD would prove valuable.

ACKNOWLEDGEMENTS

We are grateful to N. Papanicolaou for clarifying discussions. One of us (D.G.C.) thanks A. Polyakov for a stimulating conversation, and the Aspen Center for Physics where part of this work was initiated.

APPENDIX

We consider purely real solutions to (5.9). If we multiply (5.9a) by (5.9b) we obtain

$$\Phi_q e^{\Phi_q} = \Phi_p e^{\Phi_p} , \tag{A.1}$$

for all values of k. Hence there are two possibilities:

(a) $\Phi_q = \Phi_p$,

and

(b) $\Phi_q \neq \Phi_p$.

Note that for real Φ, one cannot have $\Phi_q = -\Phi_p$ since $e^x \neq -e^{-x}$.

Case (a) $\Phi_q = \Phi_p$. We wish to determine for what values of the parameter k a solution can exist to the equation

$$e^x = \pm \frac{x}{k} . \tag{A.2}$$

For the minus sign, it is easily seen that only one solution exists for any k, and that x is negative.

For the plus sign, it is also easy to show that for $k > e^{-1}$, no solution exists; for $k = e^{-1}$, the only solution is $x = 1$; and for $k < e^{-1}$, two solutions exist and they are positive.

Case (b) $\Phi_q \neq \Phi_p$. For this to be true, the function in (A.1), xe^x, must have the same value for two different values of x. This can only happen for negative x, so both Φ_q and Φ_p must be negative. By substituting (5.9b) in (5.9a) we see that Φ_q must satisfy,

$$\Phi_q = -ke^{-ke^{\Phi_q}} , \tag{A.3}$$

and only the minus sign is allowed since both functions must be negative. Let $f(x) = -ke^{-ke^x}$. As $x \to -\infty$, $f \to -k$. Also $f(0) = -ke^{-k}$. Hence, there is always one negative solution for any $k > 0$. However, since (A.3) must be satisfied for any solution, and we already know that there exists a solution for the minus sign case, namely $\Phi_q = \Phi_p = s$ and $s < 0$ for all k, the question is whether at least two other solutions exist for some k, namely $\Phi_q = d_1$, $\Phi_p = d_2$, and $\Phi_q = d_2$, $\Phi_p = d_1$ for $d_1 \neq d_2$ and both negative.

To answer this question, consider (A.3) as the function

$$F = x + ke^{-y} \tag{A.4}$$

with

$$y = ke^x \geqslant 0 . \tag{A.5}$$

It is clear that

$$F = 0 \Rightarrow x < 0 \Rightarrow y < k . \tag{A.6}$$

Furthermore, since $F(-\infty) < 0$ and $F(0) > 0$, the number of zeros of F is odd. If we assume that F has two additional solutions besides the known one, then F' has at least two different zeros, and therefore F" has a zero. Let

$$F'' (x^*) = 0, \tag{A.7a}$$

then

$$F' (x^*) < 0 . \tag{A.7b}$$

Since

$$F' = 1 - kye^{-y} , \tag{A.8}$$

$$F'' = kye^{-y} (y - 1) , \tag{A.9}$$

we have

$$F'' = 0 \Rightarrow y^* = 1 \Rightarrow x^* = -\ln k. \tag{A.10}$$

But then

$$F' (x^*) = 1 - \frac{k}{e} < 0 , \tag{A.11}$$

the inequality coming from (A.7). Hence we conclude that for $k < e$, since this does not satisfy (A.11), there is only one zero for F. Note that from (A.9) F" has only one zero, so the choice is only between one or three solutions.

For $k = e$, it is easy to check that F has a threefold zero at $x^* = -1$ $(y^* = 1)$. Indeeed $F = F' = F'' = 0$, and $F''' = 1$.

In the case $k > e$, we give a perturbative analysis around $k = e (1 + \varepsilon)$ which shows that three different solutions exist at least for small ε. Explicit calculation has convinced us that this is actually true for all $k > e$.

We want to show that F defined in (A.4) and (A.5) has three zeros for $k = e(1 + \varepsilon)$, $\varepsilon \ll 1$. Note that for $k^* = e$, $y^* = 1$. We will use expansions of F, F', and F" to first order in ε to make our argument. So we need

$$\frac{dF}{dk} = \mathring{F} = e^{-y} (1 - y) , \tag{A.12}$$

For $y^* = 1$:

$$\mathring{F} (y^*) = 0 . \tag{A.13}$$

Then

$$\mathring{F}' = y e^{-y} (y - 2) ; \mathring{F}' (y^*) = -\frac{1}{e} , \tag{A.14}$$

$$\mathring{F}'' = y e^{-y} (4y - y^2 - 2) ; \mathring{F}'' (y^*) = \frac{1}{e} . \tag{A.15}$$

So we obtain the following expansions for $k^* = e$, $y^* = 1$, and $x^* = -1$:

$$F(x^*,k) \approx F(k^*) + \mathring{F}(k^*) (k - k^*) + O(\varepsilon^2)$$

$$= 0 + \varepsilon e \mathring{F}(y^*) + O(\varepsilon^2)$$

$$= 0 + O(\varepsilon^2) , \tag{A.16}$$

$$F(x^*,k) \approx 0 + \varepsilon e \mathring{F}'(y^*) + O(\varepsilon^2)$$

$$= -\varepsilon + O(\varepsilon^2) , \tag{A.17}$$

$$F''(x^*,k) \approx 0 + \varepsilon e \mathring{F}''(y^*) + O(\varepsilon^2)$$

$$= \varepsilon + O(\varepsilon^2) . \tag{A.18}$$

Hence for $k = e(1 + \varepsilon)$, two of the zeros of F have moved to different points from the triple zero point $x^* = -1$ for $k^* = e$, since now F changes sign at x^* as follows (all $x < 0$):

$$x_1 = \tilde{x} + \delta > x > x_2 = \tilde{x} - \delta,$$

$$F(x_1) < 0 \ , \quad F(x_2) > 0 \ . \tag{A.19}$$

So at least for k which is ε bigger than e, there are exactly three solutions to (5.9). There are only three, since as noted from (A.9) F'' is zero only once.

REFERENCES

[1] For a review see S. Coleman, "Classical Lumps and Their Quantum Descendants," Int. School of Subnuclear Physics, Erice, Italy, 1975 (Plenum Press, N.Y. 1977).

[2] For a review see S. Coleman, "The Uses of Instantons," Erice Lectures, 1977.

[3] G. 't Hooft, Nucl. Phys. B 72 (1974) 461; B75 (1974) 461.

[4] E. Witten, Nucl. Phys. B160 (1979) 57.

[5] E. Witten, Nucl. Phys. B149 (1979) 285.

[6] R. J. Cant and A. C. Davis, Phys. Lett. 87B (1979) 215.

[7] H. Eichenherr, Nucl. Phys. B146 (1978) 215; B155 (1979) 544.

[8] V. L. Golo and A. M. Perelomov, Phys. Lett. 79B (1978) 112.

[9] E. Cremmer and J. Scherk, Phys. Lett. 74B (1978) 341; A. J. Mac Farlane, Nucl. Phys. B152 (1979) 145.

[10] A. D'Adda, M. Lüscher and P. Di Vecchia, Nucl. Phys. B146 (1978) 63; B152 (1979) 125.

[11] M. Lüscher, Phys. Lett. 78B (1978) 465.

[12] B. Berg and M. Lüscher, Comm. Math. Phys. 69 (1979) 57.

[13] V. A. Fateev, I.V. Frolov and A. S. Schwarz, Yad. Fiz. 30 (1979) 1134.

[14] A. M. Din, P. Di Vecchia and W.J. Zakrzewski, Nucl. Phys. B155 (1979) 447.

[15] M. Stone, Nucl. Phys. B152 (1979) 97; E. Gava, R. Jengo and C. Omero, Nucl. Phys. B158 (1979) 381; I. Ya. Aref'eva and S.I. Azakov, Nucl. Phys. B162 (1980) 298.

[16] H. E. Haber, I. Hinchliffe and E. Rabinovici, "The CP^{N-1} Model with Unconstrained Variables," LBL preprint-10519 (1980).

[17] J. Affleck "The Role of Instantons in Scale Invariant Gauge Theories," Harvard preprint - HUTP-79/A047 (July, 1979).

[18] For reviews see J. B. Kogut, Rev. Mod. Phys. 51 (1979) 659;

D. Maison, "Some Facts About Classical Non-Linear σ-Models," Max-Planck Institute preprint MPI-PAE/PTh 52/79 (Nov. 1979);

see also W. A. Bardeen, B.W. Lee and R. E. Shrock, Phys.Rev. D14 (1976) 985;

E. Brézin, J. Zinn-Justin and J.C. Le Guillou, Phys. Rev. D14 (1976) 2615.

[19] A. M. Polyakov, Phys. Lett. 59B (1975) 79;

A. A. Belavin and A. M. Polyakov, JEPT Lett. 22 (1975) 245.

[20] D. J. Gross, Nucl. Phys. B132 (1978) 439.

[21] A. Jevicki, Nucl. Phys. B127 (1977) 125.

[22] D. Förster, Phys. Lett. 66B (1977) 279; Nucl. Phys. B130 (1977) 38.

[23] V. A. Fateev, I.V. Frolov and A. S. Schwarz, Nucl. Phys. B154 (1979) 1.

[24] G. Lazarides, "Mass Generation in the O(3) Non-Linear σ Model," CERN preprint-TH.2777 (Nov. 1979).

[25] A. B. Zamolodchikov and A. B. Zamolodchikov, Nucl. Phys. B133 (1978) 525.

[26] D. Friedan, U.C. Berkeley Ph.D. thesis (1980), unpublished.

[27] A. Jevicki, Phys. Rev. D20 (1979) 3331.

[28] C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Rev. D16 (1977) 2526;

S. Raby and A. Ukawa, Phys. Rev. D18 (1978) 1154.

[29] B. Berg and M. Lüscher, Nucl. Phys. B160 (1979) 281.

[30] B. Berg, "Recent Results on the Instanton Approximation," DESY preprint-79/56 (Sept. 1979).

[31] C. G. Callan, R. F. Dashen and D. J. Gross, Phys. Rev. D20 (1979) 3279; Phys. Rev. Lett. 44 (1980) 435.

[32] J. Koplik, A. Neveu and S. Nussinov, Nucl. Phys. B123 (1977) 109.

[33] C. Bernard, Phys. Rev. D19 (1979) 3013.

[34] K. Bardakci and E. Rabinovici, Phys. Rev. D20 (1979) 1360.

[35] Z. F. Ezawa, "Topological Excitations and Quark Confinement," Max-Planck Institute preprint-MPI-PAE/PTh 40/79 (Aug. 1979).

[36] J. M. Kosterlitz and D. J. Thouless, J. Phys. C 6 (1973) 1181;

J. M. Kosterlitz, J. Phys. C 7 (1974) 1046;

J. Fröhlich, Comm. Math. Phys. 47 (1976) 233;

see also Kogut, ref. [18];

R. Savit, "Duality in Field Theory and Statistical Systems," Michigan preprint-UMHE 79/8 (1979), to be published in Rev. Mod. Phys. (April, 1980).

[37] H. Lehmann and K. Stehr, "The Bose Field Structure Associated with a Free Massive Dirac Field in One Space Dimension," DESY Preprint - 76/29 (June 1976)

[38] B. Schroer and T. Truong, Phys. Rev. D15 (1977) 1684.

[39] S. Coleman, Phys. Rev. D11 (1975) 2088.

[40] G. Woo, J. Math. Phys. 18 (1977) 1264.

[41] J. L. Gervais, A. Neveu and M. A. Virasoro, Nucl. Phys. B138 (1978) 45.

[42] For reviews see Kogut [18] and Savit [36].

[43] A. M. Din and W. J. Zakrzewski, "Embeddings of Classical Solutions of O(2p+1) Non-Linear σ Models in CP^{n-1} Models," CERN preprint-TH-2722 (Aug. 1979);

A. M. Din and W. J. Zakrzewski, "Stability Properties of Classical Solutions to Non-Linear σ Models," CERN preprint - TH-2721 (Aug. 1979).

[44] A. Jevicki, "Collective Behavior of Instantons in QCD," IAS preprint (June, 1979).

[45] For a review see S. Samuel, Phys. Rev. D18 (1978) 1916.

[46] F. W. Wiegel, Phys. Rep. C16 (1975) 57.

[47] A. M. Polyakov, Nucl. Phys. B120 (1977) 429.

[48] A. A. Belavin, V.A. Fateev, A.S. Schwarz and Yu. S. Tyupkin, Phys.Lett. 83B (1979) 317.

[49] G. 't Hooft, Nucl. Phys. B138 (1978) 1; B153 (1979) 141;

S. Mandelstam, Phys. Rev. D19 (1979) 2391;

T. Banks and A. Casher, "Chiral Symmetry Breaking in Confining Theories," Tel-Aviv preprint - TAUP 785-79 (1979).

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720