Theta-vacuum and large N limit in $\mathbb{C}P^{N-1}$ σ models

M. Aguado [a,*], M. Asorey [b]

[a] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*
[b] *Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, E-50009 Zaragoza, Spain*

Abstract

The θ dependence of the vacuum energy density in $\mathbb{C}P^{N-1}$ models is re-analysed in the semiclassical approach, the $1/N$ expansion and arguments based on the nodal structure of vacuum wavefunctionals. The $1/N$ expansion is shown not to be in contradiction with instanton physics at finite (spacetime) volume V. The interplay of large volume V and large N parameter gives rise to two regimes with different θ dependence, one behaving as a dilute instanton gas and the other dominated by the traditional large N picture, where instantons reappear as resonances of the one-loop effective action, even in the absence of regular instantonic solutions. The realms of the two regimes are given in terms of the mass gap m by $m^2V \ll N$ and $m^2V \gg N$, respectively. The small volume regime $m^2V \ll N$ is relevant for physical effects associated to the physics of the boundary, like the leading rôle of edge states in the quantum Hall effect, which, however, do not play any rôle in the thermodynamic limit at large N. Depending on the order in which the limits $N \to \infty$ and $V \to \infty$ are taken, two different theories are obtained; this is the hallmark of a phase transition at $1/N = 0$.
© 2010 Elsevier B.V. All rights reserved.

Keywords: Large N limit; Theta vacuum; Instantons

1. Introduction

A lack of nonperturbative analytical methods haunts the study of the infrared behaviour of confining field theories such as QCD. The main tools used for this purpose rely on approximations (e.g., semiclassical, large number of colours), and rigorous results are attainable in few corners of parameter space. The bordering region between topology and field physics is

---

* Corresponding author.
  E-mail addresses: miguel.aguado@mpq.mpg.de (M. Aguado), asorey@saturno.unizar.es (M. Asorey).
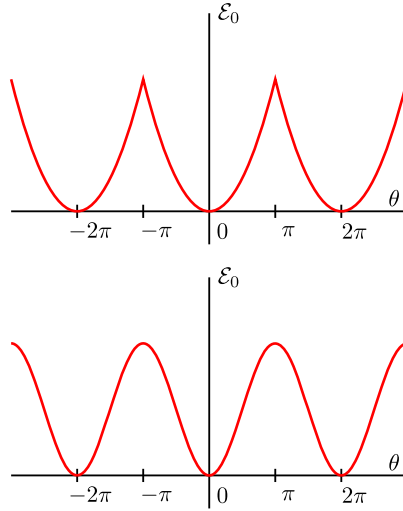
Fig. 1. Structure of the vacuum energy density in the traditional large N picture (above) and in the semiclassical picture dominated by instantons (below).

especially troubling, since different methods arise sometimes from apparently incompatible hypotheses and physical pictures.

Two-dimensional $\mathbb{CP}^{N-1}$ sigma models [1–3] are regarded as a convenient testing ground to prepare the assault on four-dimensional gauge theories, because both kinds of theories share a number of important properties: conformal invariance at the classical level, asymptotic freedom, dynamical mass generation, confinement, existence of a topological term θ and instantons for all values of the *number of colours N*.

A relevant problem in these theories with topological properties is the θ dependence of the vacuum energy density, the quantity that determines the phase structure of the theory (for a recent review, see [4]). In particular, the fate of the discrete parity symmetry P upon quantisation at the values $\theta = 0$ and $\theta = \pi$ (the only values for which it is classically conserved) is an issue.

Perhaps the simplest model in which the subtlety of the θ-dependence of the vacuum energy $\mathcal{E}_0$ is manifest is the quantum rotor [5], i.e., the quantum mechanical problem of the dynamics of a charged particle on the circumference S^1 enclosing a magnetic flux θ. In the absence of perturbations, the vacuum energy is quadratic in θ; periodicity of the physics in $\theta \to \theta + 2\pi$ imposes that the ground level is twofold degenerate for $\theta = \pi$ (i.e., half a flux quantum across the region bounded by S^1) and there parity is spontaneously broken. This, as we will see, mimics the traditional picture of the large N expansion in $\mathbb{CP}^{N-1}$ models. However, even slight perturbations compatible with reflection symmetry lift the degeneracy of the rotor, by a level repulsion mechanism, making the curve $\mathcal{E}_0(\theta)$ smooth at $\theta = \pi$. A convenient approximate method is that of the dilute instanton gas, where the vacuum is understood in terms of tunnelling processes among classical vacua. In the dilute approximation, the vacuum energy (a pure nonperturbative effect) is a smooth periodic function of θ proportional to $(1 - \cos\theta)$. This corresponds to the semiclassical approximation in $\mathbb{CP}^{N-1}$ models, where instantons play an all-important rôle.

These two regimes have the following paradigmatic expressions for $\mathcal{E}_0(\theta)$, illustrated in Fig. 1:

$$\mathcal{E}_0(\theta) \propto \min\{(\theta + 2\pi k)^2; k \in \mathbb{Z}\} \quad \text{(large } N\text{)}, \tag{1}$$

$$\mathcal{E}_0(\theta) \propto 1 - \cos\theta \quad \text{(semiclassical)}. \tag{2}$$

We now consider the situation in $\mathbb{C}P^{N-1}$ models. Exact solutions are known for the quantum $\mathbb{C}P^1$ model (equivalent to the O(3) model) both at $\theta = 0$ and $\theta = \pi$. In the first case [6], the solution exhibits a mass gap, the spectrum consists of an SU(2) triplet, and parity is conserved. This agrees with the Haldane map [7], which transforms this model into a chain of integer classical spins. Vafa and Witten [8] argued that there is no first order phase transition with spontaneous parity breaking at $\theta = 0$ for QCD, their argument being applicable straightforwardly to all $\mathbb{C}P^{N-1}$ models (see [9] for a proof of the Vafa–Witten theorem using the topological charge as an order parameter).

The exact solution of the quantum $\mathbb{C}P^1$ model at $\theta = \pi$ [10] also conserves parity but shows no mass gap (this result was anticipated in [11]). The critical behaviour of the model is described by an SU(2) WZNW model at level $k = 1$. This also agrees with the Haldane map, which transforms this model into a chain of half-odd spins. By the Lieb–Schulz–Mattis theorem [12], the absence of mass gap implies that P is conserved. On the other hand, the absence of a first order phase transition with spontaneous P breakdown at $\theta = \pi$ has been argued to hold for all $\mathbb{C}P^{N-1}$ models [13], by analyzing the nodal structure of the vacuum in the Hamiltonian formalism [14] in analogy with QCD [15].

For the intermediate region $0 < \theta < \pi$, analytical techniques are lacking, and we must rely on approximations and numerical simulations. We will discuss two important approximations, which have been argued to be mutually incompatible: the semiclassical method and the $1/N$ expansion.

The semiclassical approach [16] is based, as in the case of the rotor, on the picture of the quantum vacuum of $4d$ gauge theories and $2d$ $\mathbb{C}P^{N-1}$ models built from tunnelling processes among classical vacua. These nonperturbative processes are dominated by instantons and antiinstantons, (anti)selfdual solutions of the classical Euclidean equations of motion. A dilute gas approximation gives a θ dependence of the vacuum energy density of the form

$$\mathcal{E}_0(\theta) \propto m^2(1 - \cos\theta), \tag{3}$$

where m is the mass gap. This dependence cannot be seen in perturbation theory due to the nonanalytic dependence of the mass gap on the coupling.

However, a vacuum based on a dilute gas of instantons and antiinstantons is not satisfactory, since the statistical ensemble is dominated by the infrared divergent contribution of arbitrarily large instantons, whose density n as a function of size ρ is

$$n(\rho)\mathrm{d}\rho \propto (\Lambda\rho)^N \frac{\mathrm{d}\rho}{\rho^3} \tag{4}$$

for the $\mathbb{C}P^{N-1}$ model, with Λ a typical scale of the theory. A statistical mechanical treatment of *interacting* instanton fluids has been developed [17,18], bringing about the instanton liquid picture of the QCD vacuum [19]. This may be very relevant for the behaviour of these theories at finite temperature and high density. We note in passing that the dilute gas approximation breaks down as well in the *ultraviolet* for the $\mathbb{C}P^1$ model, as pointed out by Lüscher [20] building on his work with Berg [21] on the geometric definition of a topological charge density on the lattice. Technically, the topological susceptibility in this model does not scale according to the perturbative renormalisation group due to small distance fluctuations. This is reflected in the singularity of (4) as $\rho \to 0$ for $N = 2$; it may also be understood a consequence of the slow vanishing rate of the density of Lee–Yang zeros as $\theta \to 0$ [9]. What is remarkable is that, although the semiclassical analysis does not reveal any ultraviolet instanton singularity in the case of $\mathbb{C}P^2$ model, numerical simulations suggest a similar pathology [22].

The $1/N$ expansion [23,24] stands as an alternative to the semiclassical method. This technique is based on the simplification of both $4d$ SU(N) gauge theories and $2d$ $\mathbb{C}P^{N-1}$ models when N is taken to infinity keeping certain parameter combinations constant.

The $1/N$ expansion of $\mathbb{C}P^{N-1}$ models, as developed in [3] and [25], agrees at $\theta = 0$ with the known spectrum, given by a massive particle in the adjoint representation of SU(N). The mass m_T is generated dynamically, and the particle turns out to be a composite state of two fundamental fields, bound together by a Coulomb potential. For $\theta \neq 0$, this analysis predicts a quadratic θ dependence

$$\mathcal{E}_0(\theta) = \frac{3}{2\pi} \frac{m_T^2 \theta^2}{N} \tag{5}$$

of the vacuum energy density around $\theta = 0$. This dependence can be made to agree with the fundamental requirement that physics be periodic in θ with period 2π only if there is a first order cusp at odd multiples of $\theta = \pi$, i.e., a first order phase transition accompanied by spontaneous parity breakdown, as shown in the upper part of Fig. 1.

Instanton effects, being nonperturbative, are not visible in the perturbative expression (5). This led Witten [25] to argue that the $1/N$ expansion is not sensitive to instantons — equivalently, that instantons play no significant rôle in the quantum $\mathbb{C}P^{N-1}$ models (or in $4d$ gauge theories) to the extent that the $1/N$ expansion is a good approximation thereof. Jevicki [26], however, argued that instantons resurface in the $1/N$ expansion as *poles* of the integrand of the partition function Z (Eq. (8)), and that Z can be computed both by the saddle point method and by using a functional Cauchy theorem summing the residues of all these poles (representing resonances). Then the large N limit and instanton effects would not be *a priori* incompatible with each other.

The quadratic dependence (5) agrees with the holographic picture provided by the Maldacena conjecture [27], and moreover with lattice measurements of the topological susceptibility of $\mathbb{C}P^{N-1}$ models (see [4] for a review). The Witten–Veneziano formula [28,29], derived in this approximation, gives a phenomenologically correct value of the η' mass in terms of the topological susceptibility at $\theta = 0$. However, the appearance of a first order cusp at $\theta = \pi$ is in contradiction with the results arising from the nodal analysis of the vacuum [13], and with the intuition that level repulsion generically destroys level crossings.

In this work we show how this discrepancy stems from the fact that the large N limit and the thermodynamical limit do not commute. The traditional formulation of the $1/N$ expansion starts directly at infinite spacetime volume $V = LT = \infty$. As we shall see, a procedure in which the thermodynamic limit is taken after the $N \to \infty$ limit provides results compatible both with instanton physics and with the rigorous results at $\theta = \pi$, and different from the reverse order of limits. A finite volume analysis is in order. This agrees with Schwab's [30,31] and Münster's [32,33] approach in the case of the sphere; we moreover outline the application of Jevicki's residue method. We find the case of the torus much more tractable and amenable to explicit computation after integration over the dual torus parametrising the different holomorphic bundle structures within each topological charge sector. In particular, Jevicki's approach requires computation of the residues of meromorphic functions instead of functionals, and his programme can be carried out in the simplest cases exhibiting the reappearance of instantons as resonances (poles) in the one-loop effective action. Remarkably, in spite of the absence of regular unit charge instantons on the torus [34], the contribution of this sector is nonzero in the resonance approach: this we interpret as the effect of rough configurations near the forbidden regular instanton.

At finite volume there are two regimes: one dominated by instantons for low mass theories, $m^2 V \ll N$, and another regime where they are strongly suppressed, $m^2 V \gg N$. The second regime is the relevant one for $\mathbb{CP}^{N-1}$ theories in the thermodynamic limit, but the other regime is relevant for effects where the finite volume or space topology play a leading rôle, like in the appearance of edge states in the quantum Hall effect.

The structure of the article is as follows. In Section 2, the traditional large N picture of $\mathbb{CP}^{N-1}$ models is reviewed. The θ dependence of $\mathbb{CP}^{N-1}$ models formulated on the sphere is considered in Section 3. The corresponding analysis for the case of the torus is performed in Section 4. The consequences of this analysis are discussed in Section 5.

2. The traditional picture of the $1/N$ expansion

The traditional large N picture of $\mathbb{CP}^{N-1}$ models was developed in [3] and [25]. We shall now give a brief account of it before the analysis in finite volume.

The large N method is based in a saddle point approximation of the partition function, defined on the infinite $2d$ Euclidean plane, after integration of the fundamental Ψ, $\Psi^\dagger$ fields (taking values in $\mathbb{C}^N$, i.e., in representatives of projective classes in $\mathbb{CP}^{N-1}$). We introduce the dummy U(1) gauge field

$$A_\nu = -\frac{i}{2}\big(\Psi^\dagger \partial_\mu \Psi - (\partial_\mu \Psi)^\dagger \Psi\big), \tag{6}$$

and a scalar field $\alpha(x)$ imposing the constraint $\Psi^\dagger \Psi = 1$ at each point as a Lagrange multiplier. Starting from the full partition function at $\theta = 0$,

$$Z = \int \mathcal{D}\Psi \mathcal{D}\Psi^\dagger \mathcal{D}A_\mu \, \delta\big[\Psi^\dagger \Psi - 1\big] \exp\bigg\{-\frac{N}{2g_0^2} \int_{\mathbb{R}^2} d^2x \, |D_\mu \Psi|^2\bigg\}$$

$$= \int \mathcal{D}\Psi \mathcal{D}\Psi^\dagger \mathcal{D}A_\mu \mathcal{D}\alpha \exp\bigg\{-\frac{N}{2g_0^2} \int_{\mathbb{R}^2} d^2x \, |D_\mu \Psi|^2 - \frac{N}{2g_0^2} \int_{\mathbb{R}^2} d^2x \, \alpha(x)\big(\Psi^\dagger \Psi - 1\big)\bigg\}, \tag{7}$$

we perform the Gaussian integration over Ψ, $\Psi^\dagger$ to obtain

$$Z = \int \mathcal{D}A_\mu \mathcal{D}\alpha \, e^{-N S_{\mathrm{eff}}[A_\mu, \alpha]}, \tag{8}$$

where the effective action is

$$S_{\mathrm{eff}}[A_\mu, \alpha] = \mathrm{Tr} \ln\big(-D_\mu^2 - \alpha(x)\big) - \frac{1}{2g_0^2} \int_{\mathbb{R}^2} d^2x \, \alpha(x). \tag{9}$$

The saddle point equations

$$\frac{\delta S_{\mathrm{eff}}}{\delta \alpha(x)} = \frac{1}{-D_\mu^2 + \alpha}(x, x) - \frac{1}{2g_0^2} = 0, \tag{10}$$

$$\frac{\delta S_{\mathrm{eff}}}{\delta A_\mu(x)} = 2i \frac{D_\mu}{-D_\mu^2 + \alpha}(x, x) = 0, \tag{11}$$

can be solved within a renormalisation scheme to yield a saddle configuration

$$A_\mu = 0, \qquad \alpha = m_{\mathrm{T}}^2 \equiv \mu^2 \exp\bigg\{-\frac{2\pi}{g_R^2(\mu)}\bigg\}, \tag{12}$$

where μ is a mass scale and g_R the corresponding renormalised coupling.

Perturbation theory around the saddle configuration reveals a dynamical system of an N-plet of charged scalars Ψ, with a short range interaction due to the field α, and electromagnetic interaction due to the field A_μ. The latter develops an effective kinetic term and couples to the scalars with effective electric charge $e_{\text{eff}} = \sqrt{12\pi m_T^2/N}$. Thus there is a confining Coulomb interaction between scalars, and the spectrum at $\theta = 0$ consists of $(\Psi\Psi^\dagger)$ bound states, with mass gap $2m_T$, living in the adjoint representation of SU(N).

For $\theta \neq 0$, the topological term in the action is

$$-i\theta Q = -i\frac{\theta}{2\pi}\int d^2x\, F_{01}, \tag{13}$$

Q being the magnetic flux associated with the U(1) field A_μ and its field strength $F_{\mu\nu}$, or equivalently, the topological charge of the field Ψ, which is an integer for smooth finite-action configurations. This term plays the rôle of an external electric field in electrodynamics. Therefore, in this picture of the $\mathbb{CP}^{N-1}$ models, its contribution to the vacuum energy density is

$$\mathcal{E}_0(\theta) = \frac{1}{2}e_{\text{eff}}^2\left(\frac{\theta}{2\pi}\right)^2 = \frac{3}{2\pi}\frac{m_T^2\theta^2}{N}, \tag{14}$$

yielding a topological susceptibility

$$\chi_t = \left(\frac{d^2\mathcal{E}_0(\theta)}{d\theta^2}\right)_{\theta=0} = \frac{3m_T^2}{\pi N}. \tag{15}$$

This quadratic dependence is perturbative, i.e., it can be seen in terms of Feynman diagrams. Instanton effects, nonperturbative in nature, were argued in [25] to be exponentially suppressed in the $1/N$ expansion, and therefore irrelevant for the physics of the $\mathbb{CP}^{N-1}$ models. However, we have seen that the level crossing and first order phase transition at $\theta = \pi$ implied by (14) and the requirement of 2π-periodicity in θ are in contradiction with the nodal arguments of [13]. We will next go over to a compact space with the purpose of showing that this incompatibility stems from the infinite volume starting point of the traditional $1/N$ analysis.

3. $1/N$ expansion on S^2

Clarifying the interplay of N and the volume, and the effects of taking these to infinity in different orders, requires the $\mathbb{CP}^{N-1}$ models to be first formulated in a compact (Euclidean) space.

Schwab [30,31] and Münster [32,33] studied the $1/N$ expansion of $\mathbb{CP}^{N-1}$ models on S^2, and observed that the $k = 1$ contribution to the partition function is dominated, for large N, by a saddle point given by a rotationally invariant instanton (in the sense that global U(N) transformations can be compensated by O(3) rotations in Euclidean space). The saddle point equations

$$\frac{1}{-D^\mu D_\mu + \alpha}(x,x) = \frac{1}{2g_0^2},$$
$$\frac{D_\nu}{-D^\mu D_\mu + \alpha}(x,x) = 0, \tag{16}$$

admit for $N > |k|$, in a uniform topological charge density background, solutions with constant $\alpha(x)$. Indeed, the second equation holds due to parity. The first equation states rotation invariance

of the propagator $G(x, x)$ of a particle with mass $\sqrt{\alpha}$. It is easy to show that $G(x, y)$ depends only on the geodesic distance between x and y, and therefore the first equation has solutions with constant α.

3.1. The effective action on the sphere

Thus, we begin [35,36] with a spherical spacetime of radius R and volume $V = 4\pi R^2$, and the action of the $\mathbb{C}P^{N-1}$ model on a background of topological charge k,

$$S_k = -\frac{N}{2g_0^2} \int \Psi^\dagger \Delta_k \Psi + \frac{N}{2g_0^2} \int m^2 (\Psi^\dagger \Psi - 1), \tag{17}$$

where integration implies the measure $\mathrm{d}^2 x \sqrt{g}$, and Δ_k is the covariant Laplacian in the background chosen. The magnetic flux for the composite U(1) field is quantised,

$$\Phi_B = 4\pi R^2 B = 2\pi k, \quad k \in \mathbb{Z}. \tag{18}$$

We rewrite the constant saddle point value of the α field as m^2, variable still to be integrated upon.

Integrating out Ψ, $\Psi^\dagger$ yields the functional determinant of the operator $-\Delta_k + m^2$, which is computed in the ζ function renormalisation scheme at energy scale μ. Discarding unessential factors,

$$Z_k = \int \mathrm{d}m^2 \, \mathrm{e}^{-N S_k^{\mathrm{eff}}}, \tag{19}$$

with effective action

$$S_k^{\mathrm{eff}} = \ln \det \left(\frac{-1}{\mu^2} \Delta_k + \frac{m^2}{\mu^2} \right) - \frac{4\pi R^2}{2g_0^2} m^2$$

$$\equiv \ln \det \mathcal{A} - \frac{4\pi R^2}{2g_0^2} m^2. \tag{20}$$

The eigenvalues of $\mathcal{A}$ are

$$\lambda_n = \frac{1}{\mu^2 R^2} \left[\left(n + \frac{|k|}{2} + 1 \right) \left(n + \frac{|k|}{2} \right) - \frac{k^2}{4} + m^2 R^2 \right]$$

$$\equiv \frac{\widetilde{\lambda}_n}{\mu^2 R^2}, \quad n = 0, 1, 2, \dots, \tag{21}$$

with degeneracy $d_n = 2n + |k| + 1$.

We use the ζ-function definition of the determinant, equivalent to a renormalisation at scale μ:

$$\ln \det_\zeta \mathcal{A} = \sum_{n=0}^{\infty} d_n \ln \lambda_n = \sum_{n=0}^{\infty} d_n \ln \widetilde{\lambda}_n - \left(\sum_{n=0}^{\infty} d_n \right) \ln \left(\mu^2 R^2 \right)$$

$$\to -\zeta'_{\widetilde{\mathcal{A}}}(0) - \zeta_{\widetilde{\mathcal{A}}}(0) \ln \left(\mu^2 R^2 \right), \tag{22}$$

$\zeta_{\widetilde{\mathcal{A}}}$ being the ζ function associated with the operator $\widetilde{\mathcal{A}} \equiv \mu^2 R^2 \mathcal{A}$, i.e. the analytic continuation of

$$\zeta_{\widetilde{\mathcal{A}}}(s) \equiv \sum_{n=0}^{\infty} \frac{d_n}{\widetilde{\lambda}_n^s}, \quad (\mathrm{Re}\, s > 1) \tag{23}$$
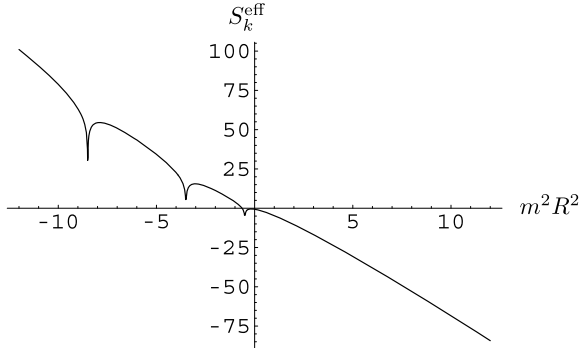
Fig. 2. Effective action in the $k = 1$ sector on the sphere, for real $m^2 R^2$.

for all complex $s \neq 1$.

From the small s expansion of (23), we obtain the effective action for topological sector k:

$$
\begin{aligned}
S_k^{\text{eff}} &= -\zeta'_{\tilde{\mathcal{A}}}(0) - \zeta_{\tilde{\mathcal{A}}}(0) \ln(\mu^2 R^2) - \frac{4\pi R^2}{2 g_R^2} m^2 \\
&= 2\left(\frac{k^2}{4} + \frac{1}{4} - m^2 R^2 \right) + \left(m^2 R^2 - \frac{1}{3} \right) \ln(\mu^2 R^2) - \frac{4\pi R^2}{2 g_R^2} m^2 \\
&\quad + 2\sqrt{\frac{k^2}{4} + \frac{1}{4} - m^2 R^2} \, \ln \frac{\Gamma(\frac{|k|+1}{2} + \sqrt{\frac{k^2}{4} + \frac{1}{4} - m^2 R^2})}{\Gamma(\frac{|k|+1}{2} - \sqrt{\frac{k^2}{4} + \frac{1}{4} - m^2 R^2})} \\
&\quad - 2\zeta'_H\left(-1; \frac{|k|+1}{2} + \sqrt{\frac{k^2}{4} + \frac{1}{4} - m^2 R^2} \right) \\
&\quad - 2\zeta'_H\left(-1; \frac{|k|+1}{2} - \sqrt{\frac{k^2}{4} + \frac{1}{4} - m^2 R^2} \right).
\end{aligned}
\tag{24}
$$

Here we have used the Hurwitz zeta function $\zeta_H(s; v)$, defined by analytical continuation to all $s \neq 1$ of

$$
\zeta_H(s; v) = \sum_{n=0}^{\infty} (n + v)^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty dt \, \frac{t^{s-1} e^{-vt}}{1 - e^{-t}}, \quad \operatorname{Re} s > 1,
\tag{25}
$$

and its derivative $\zeta'_H(s; v)$ with respect to s. Function (24) is defined for all complex values of $m^2 R^2$, bar isolated singularities.

3.2. Zeros and saddle points of the effective action

In order to compute Z_k, integration over m^2 is still to be performed, through imaginary values in order to ensure convergence.

Let us study the behaviour of the effective action for *real* $m^2 R^2$ (see Fig. 2). To begin with, the integrand of Z_k has $N(2n + |k| + 1)$-fold poles at

$$m^2 R^2 = p_n = -\left(n + \frac{|k|}{2} + 1\right)\left(n + \frac{|k|}{2}\right) + \frac{k^2}{4}, \quad n = 0, 1, 2, \ldots, \tag{26}$$

reproducing the eigenvalues and degeneracies of $-\Delta_k + m^2$. In this sense, in the way predicted by Jevicki, the partition function can be computed by deforming the integration curve so as to surround the poles, and summing the residues at each of them. However, the problem of computing and summing the residues is too difficult to be tackled analytically, although the previous formulae can be used in a numerical approach (more progress can be made analytically in the case of the torus, as will be seen in next section).

Alternatively, we can use the saddle point method. The zeros of the derivative,

$$\frac{dS_k^{\text{eff}}}{d(m^2 R^2)} = \ln(\mu^2 R^2) - \frac{4\pi}{2g_R^2}$$

$$- \psi\left(\frac{|k| + 1}{2} + \sqrt{\frac{k^2 + 1}{4} - m^2 R^2}\right)$$

$$- \psi\left(\frac{|k| + 1}{2} - \sqrt{\frac{k^2 + 1}{4} - m^2 R^2}\right), \tag{27}$$

alternate with the poles in the real $m^2 R^2$ axis, as seen in Fig. 2 (here $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function). There is a unique saddle point s_0 to the right of the first pole $p_0 = -\frac{|k|}{2}$, which we assume to be dominant.

The partition function of sector k in the saddle point approximation is

$$Z_k^{(s_0)} = \frac{1}{R^2} e^{-N S_k^{\text{eff}}(s_0)} \sqrt{\frac{2\pi}{N|S_k^{\text{eff}\prime\prime}(s_0)|}}, \tag{28}$$

up to quadratic order. Explicit results can be obtained for large $m^2 R^2$, in which region the effective action can be expanded as

$$S_k^{\text{eff}} = -\left(m^2 R^2 - \frac{1}{3}\right)\frac{m^2}{\mu^2} - \frac{4\pi R^2}{2g_R^2}m^2 + m^2 R^2$$

$$+ \left(\frac{k^2}{24} - \frac{1}{15}\right)\frac{1}{m^2 R^2} + \left(\frac{k^2}{40} - \frac{4}{315}\right)\frac{1}{m^4 R^4} + \mathcal{O}(m^{-6} R^{-6}), \tag{29}$$

and the saddle point is found to be

$$m_k^2 R^2 = m_T^2 R^2 + \frac{1}{3} - \left(\frac{k^2}{24} - \frac{1}{90}\right)\frac{1}{m_T^2 R^2}$$

$$- \left(\frac{k^2}{45} - \frac{16}{2835}\right)\frac{1}{m_T^4 R^4} + \mathcal{O}(m^{-6} R^{-6}), \tag{30}$$

where $m_T^2 = \mu^2 \exp\{-2\pi/g_R^2\}$ is the infinite volume saddle point in (12).

The total partition function after summing all topological sectors is

$$Z^{(s_0)}(\theta) = \sum_{k \in \mathbb{Z}} Z_k^{(s_0)} e^{-ik\theta}$$

$$= \sqrt{\frac{2\pi}{N} \frac{m_{\rm T}}{R}} \exp\left\{ \frac{2\pi N}{3g_R^2} - N m_{\rm T}^2 R^2 + \frac{N}{90 m_{\rm T}^2 R^2} \right\}$$

$$\times \sum_{k \in \mathbb{Z}} \exp\left\{ -\frac{N k^2}{24 m_{\rm T}^2 R^2} - ik\theta \right\} \left(1 + \mathcal{O}(m_{\rm T}^{-4} R^{-4})\right)$$

$$\equiv \sqrt{\frac{2\pi}{N} \frac{m_{\rm T}}{R}} \exp\left\{ \frac{2\pi N}{3g_R^2} - N m_{\rm T}^2 R^2 + \frac{N}{90 m_{\rm T}^2 R^2} \right\}$$

$$\times \vartheta_3\left(\frac{\theta}{2\pi} \,\bigg|\, \frac{iN}{24\pi m_{\rm T}^2 R^2} \right) \left(1 + \mathcal{O}(m_{\rm T}^{-4} R^{-4})\right), \tag{31}$$

where the last equation uses Jacobi's ϑ_3 function:

$$\vartheta_3(z|\tau) = \sum_{n \in \mathbb{Z}} e^{i\pi\tau n^2} e^{i2\pi n z}. \tag{32}$$

Two asymptotic regimes for (31) can be analysed. For $N \gg m_{\rm T}^2 R^2$, the sum therein can be truncated, keeping just the $k = -1, 0, 1$ sectors. Then the vacuum energy density has a typical dilute instanton gas θ dependence,

$$\mathcal{E}_0(\theta) - \mathcal{E}_0(0) = -\frac{1}{4\pi R^2} \ln \frac{Z^{(s_0)}(\theta)}{Z^{(s_0)}(0)}$$

$$\approx \frac{1}{4\pi R^2} \exp\left\{ -\frac{N}{24 m_{\rm T}^2 R^2} \right\} (1 - \cos\theta). \tag{33}$$

But if $m_{\rm T}^2 R^2 \gg N$, using the Poisson resummation formula for the θ function in (31) and keeping the dominant term in the dual sum, we have

$$Z^{(s_0)}(\theta) \approx \frac{4\sqrt{3}\pi m_{\rm T}^2}{N} \exp\left\{ \frac{2\pi N}{3g_R^2} - N m_{\rm T}^2 R^2 + \frac{N}{90 m_{\rm T}^2 R^2} - \frac{6 m_{\rm T}^2 R^2}{N} \widetilde{\theta}^2 \right\}, \tag{34}$$

$\widetilde{\theta}$ being the angle in $(-\pi, +\pi]$ differing from θ in an integer. The corresponding vacuum energy density coincides with the traditional large N prediction:

$$\mathcal{E}_0^{(s_0)}(\theta) - \mathcal{E}_0^{(s_0)}(0) \approx \frac{3 m_{\rm T}^2}{2\pi N} \widetilde{\theta}^2, \tag{35}$$

which is periodic in θ and undergoes first order phase transitions with level crossing at $\theta = (2\ell + 1)\pi$, $\ell \in \mathbb{Z}$.

Before commenting on these two different limiting procedures, let us perform the same analysis on the torus [35,36].

4. $1/N$ expansion on T^2

We consider a toric spacetime of linear size L and spacetime volume $V = L^2$. Functional integration over the fields of the $\mathbb{C}P^{N-1}$ model on the torus involves an additional variable, the complex coordinate $u \in \hat{T}^2$ in the dual torus parametrising the different holomorphic bundle structures associated with the complex line bundle $E_k(T^2, \mathbb{C})$ [37].

In this case, the effective action $S_u^{\mathrm{eff}}[A_\mu, \alpha]$ resulting from integration of the Ψ, $\Psi^\dagger$ fields does not have saddle points. Specifically, the saddle point equations

$$\frac{1}{-\mathrm{D}_\mu^2 + \alpha}(x, x) = \frac{1}{2g_0^2},$$

$$\frac{\mathrm{D}_\mu}{-\mathrm{D}_\mu^2 + \alpha}(x, x) = 0, \tag{36}$$

do not have solutions with constant α and topological charge density.

4.1. Quantum saddle points and the effective action on T^2

The arguments used for the sphere can, nevertheless, be adapted to the torus, generalising the saddle point method. By integrating over u (i.e., averaging over u), the irregularities of the saddle point configurations are swept off [35,36]. Upon integration over u, a reduced effective action S_{red} obtains,

$$\exp\{-S_{\mathrm{red}}[A_\mu, \alpha, N]\} = \int_{\hat{\mathrm{T}}^2} \mathrm{d}^2 u \, \mathrm{e}^{-N S_u^{\mathrm{eff}}[A_\mu, \alpha]}, \tag{37}$$

which can be argued to be dominated by constant topological density in the large N limit. The generalised saddle point equations

$$\frac{\delta}{\delta A_\mu} \frac{\partial}{\partial N} S_{\mathrm{red}}[A_\mu, \alpha, N] = 0,$$

$$\frac{\delta}{\delta \alpha} \frac{\partial}{\partial N} S_{\mathrm{red}}[A_\mu, \alpha, N] = 0, \tag{38}$$

hold in this case because they remain finite as $N \to \infty$. Their solutions we call *quantum saddle points*.

To compute the effective action in the sectors of nonzero topological charge, the ζ function method is used, renormalising at energy scale μ:

$$S_k^{\mathrm{eff}} = \ln \det\left(\frac{-1}{\mu^2}\Delta_k + \frac{m^2}{\mu^2}\right) - \frac{m^2 L^2}{2g_0^2}$$

$$\equiv -\zeta'_{\mathcal{B}}(0) - \zeta_{\mathcal{B}}(0) \ln \frac{\mu^2 L^2}{4\pi |k|} - \frac{m^2 L^2}{2g_0^2}. \tag{39}$$

The spectrum of the Laplacian on the torus, in a background with uniformly distributed topological charge $2\pi k \neq 0$, is independent of the holonomies u and consists of Landau levels $-2\omega(n + 1/2)$, $n = 0, 1, \ldots$, with $\omega = |B| = \frac{2\pi |k|}{L^2}$, where L is the linear size of the torus. These levels have $|k|$-fold degeneracy [38]. The zeta function for operator $\mathcal{B} = (-L^2\Delta_k + m^2 L^2)/(4\pi |k|)$ is

$$\zeta_{\mathcal{B}}(s) = \sum_{n=0}^\infty |k|\left(n + \frac{m^2 L^2}{4\pi |k|} + \frac{1}{2}\right)^{-s} = |k|\zeta_H\left(s; \frac{m^2 L^2}{4\pi |k|} + \frac{1}{2}\right)$$

$$= -|k|\frac{m^2 L^2}{4\pi |k|} + s|k|\ln\left\{\frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{m^2 L^2}{4\pi |k|} + \frac{1}{2}\right)\right\} + \mathcal{O}(s^2), \tag{40}$$

yielding the effective action (see [39])

$$
\begin{aligned}
S_k^{\text{eff}} &= -\frac{m^2 L^2}{4\pi}\left\{\frac{2\pi}{g_R^2} + \ln\frac{4\pi|k|}{\mu^2 L^2}\right\} - |k|\ln\left\{\frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{m^2 L^2}{4\pi|k|} + \frac{1}{2}\right)\right\}\\
&= -\frac{m^2 L^2}{4\pi}\ln\frac{4\pi|k|}{m_T^2 L^2} - |k|\ln\left\{\frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{m^2 L^2}{4\pi|k|} + \frac{1}{2}\right)\right\}
\end{aligned}
\tag{41}
$$

where m^2 is the (constant) saddle point value of the α field, and $g_R = g_R(\mu)$ is the renormalised coupling at scale μ. In the last equation, $m_T^2 = \mu^2 \exp\{-2\pi/g_R^2(\mu)\}$ stands for the large N dynamically generated mass at infinite volume.

Expression (41) can be checked to coincide with the dominant term in a large volume, constant B expansion of the corresponding effective action (24) for the sphere:

$$
S_{\text{ef},k}^{S^2} \xrightarrow{V\to\infty} S_{\text{ef},k}^{T^2} + \mathcal{O}(V^0) \quad (B = \text{const}).
\tag{42}
$$

The contribution of topological sector k to the partition function on the torus now depends only on m^2, all other fields having been integrated out:

$$
Z_k = \int \mathrm{d}m^2\, e^{-N S_k^{\text{ef}}}.
\tag{43}
$$

As in the previous sections, the integration is performed through imaginary values of m^2 to guarantee convergence.

In order to make the functional dependences in some the following expressions clear, it is useful to define dimensionless variables

$$
y = \frac{m^2 L^2}{4\pi|k|}, \qquad y_0 = \frac{m_T^2 L^2}{4\pi|k|}.
\tag{44}
$$

Then the k-sector partition function is

$$
Z_k = \frac{4\pi|k|}{L^2}\int \mathrm{d}y \left(\frac{\Gamma(y+\frac{1}{2})}{\sqrt{2\pi}}\right)^{N|k|} e^{-N|k|y\ln y_0} = \frac{4\pi|k|}{L^2}\int \mathrm{d}y\, e^{-N|k|\widetilde{S}(y)},
\tag{45}
$$

where the function

$$
\widetilde{S}(y) = \frac{S_k^{\text{eff}}}{|k|} = y\ln y_0 - \ln\frac{\Gamma(y+1/2)}{\sqrt{2\pi}}
\tag{46}
$$

is defined for all complex values of y, except for a series of poles of the integrand of Z_k, as can be seen in Fig. 3.

For small values of y, that is, when $|k| \ll m^2 L^2$, the exponent simplifies:

$$
\widetilde{S}(y) = -y\ln\frac{y}{y_0} + y + \frac{1}{24y} - \frac{7}{2880 y^3} + \mathcal{O}(y^{-4}),
\tag{47}
$$

meaning that the effective action has an expansion in powers of the topological number k where the first nontrivial term is quadratic:

$$
S_k^{\text{eff}} = -\frac{m^2 L^2}{4\pi}\ln\frac{m^2}{m_T^2} + \frac{m^2 L^2}{4\pi} + \frac{\pi k^2}{6 m^2 L^2} + \mathcal{O}\left(\frac{k^4}{m^6 L^6}\right).
\tag{48}
$$

In the opposite limit, when $|k| \gg m^2 L^2$, that is, for large y,

$$
\widetilde{S}(y) = \frac{\ln 2}{2} + \{\ln y_0 - \psi(1/2)\}y - \frac{\pi^2}{4}y^2 - \frac{1}{6}\psi''(1/2)y^3 - \frac{\pi^2}{24}y^4 + \mathcal{O}(y^5).
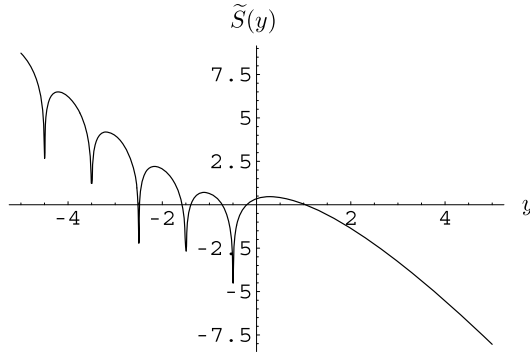\tag{49}
$$

Fig. 3. Effective action $\widetilde{S}(y)$ for the torus, with $y \in \mathbb{R}$.

Written in terms of the effective action, we see that the leading term is linear in the absolute value of the topological number:

$$S_k^{\text{eff}} = \frac{\ln 2}{2}|k| + \frac{m^2 L^2}{4\pi}\left\{ \ln \frac{m_{\mathrm{T}}^2 L^2}{4\pi |k|} - \psi(1/2) \right\} + \mathcal{O}\left(\frac{m^4 L^4}{|k|}\right). \tag{50}$$

Notice the change of asymptotic behaviour of the effective action S_k^{eff} in the different topological sectors. According to Eqs. (48) and (50), for small values of the topological charge, $|k| < m^2 V$, the effective action is quadratic in k, whereas for large topological charges its leading term is linear in $|k|$ [9]. This change of asymptotic behaviour has important physical consequences.

4.2. Zeros and saddle points on the torus

Extrema of $\widetilde{S}(y)$ allow us to perform a saddle point approximation by deforming the integration contour so that it passes through the dominant extremum. The exponent in the Z_k integral has an overall $N|k|$ factor, therefore the saddle point approximation is a large $N|k|$ expansion, and results in terms of y are general for all $k \neq 0$ sectors.

Poles of the integrand of Z_k give us a chance of testing Jevicki's proposal, since functional integration has been reduced to integration along a path in the complex plane. The integration contour must be deformed so that it surrounds each pole. In spite of the fact that various fields have been integrated out in the effective action we are working with, we shall see that instantons reappear in these poles.

The structure of saddle points of $\widetilde{S}_k$ and poles of the integrand of Z_k is represented in Fig. 4.

4.2.1. Poles and Jevicki's approach

Let us consider the poles. There is an infinite series of $N|k|$-fold poles at values m_n of m such that

$$y = \frac{m_n^2 L^2}{4\pi |k|} \equiv p_n = -\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \ldots. \tag{51}$$

Their multiplicity is equal to the complex dimension of the moduli space of charge k instantons in the $\mathbb{C}\mathrm{P}^{N-1}$ model (see [40]).

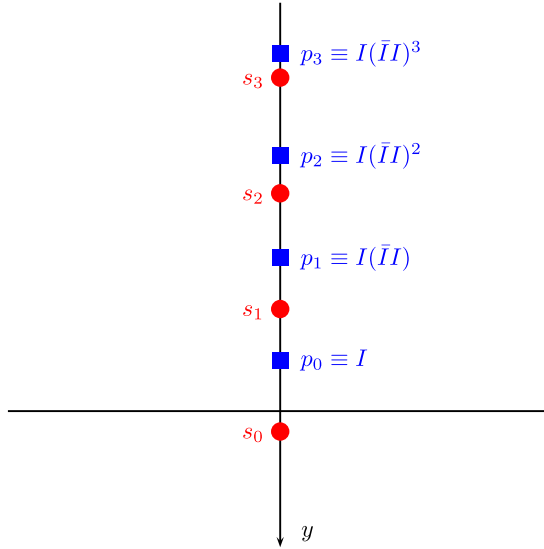The partition function can be written as a sum over residues à la Jevicki,

Fig. 4. Structure of poles p_n of the effective action (squares) and saddle points s_n of the integrand of the partition function (circles) on the torus, in the complex plane of the variable $y = m_{\mathrm{T}}^2 L^2/(4\pi|k|)$. Note that zeros and poles tend to coalesce as n grows.

$$Z_k = \frac{8\pi^2|k|}{L^2} \sum_{n=0}^{\infty} \operatorname*{Res}_{y \to p_n} \left(\frac{\Gamma(y + \frac{1}{2})}{\sqrt{2\pi}} \right)^{N|k|} e^{-N|k|y \ln y_0}$$

$$= \frac{8\pi^2|k|}{L^2} \left(\frac{y_0}{2\pi} \right)^{\frac{N|k|}{2}} \sum_{n=0}^{\infty} e^{nN|k| \ln y_0} \operatorname*{Res}_{\varepsilon \to 0} \left\{ y_0^{-\varepsilon} \Gamma(-n + \varepsilon) \right\}^{N|k|}. \tag{52}$$

There is no difficulty in computing and summing the residues for the first cases, $N|k| = 1$ and $N|k| = 2$:

$$Z_{N|k|=1} = \sqrt{2}\frac{2\pi m_{\mathrm{T}}}{L} \exp\left\{ -\frac{m_{\mathrm{T}}^2 L^2}{4\pi} \right\} \tag{53}$$

and

$$Z_{N|k|=2} = m_{\mathrm{T}}^2 \mathrm{K}_0\left(\frac{2m_{\mathrm{T}}^2 L^2}{4\pi|k|} \right) \tag{54}$$

(where K_0 is a modified Bessel function), but the partition function for higher values of $N|k|$ turns out to be more difficult to compute. From the expansion of (52), we can write it as

$$Z_k = \frac{8\pi^2|k|}{L^2} \left(\frac{y_0}{\sqrt{2\pi}} \right)^{\frac{N|k|}{2}} \sum_{n=0}^{\infty} \left(\frac{(-y_0)^n}{n!} \right)^{N|k|} T_{N|k|-1,n}(y_0). \tag{55}$$

The function $T_{R,n}(y_0)$ is given by

$$T_{R,n}(y_0) = \sum_{r,s,t=0}^{\infty} \delta_{r+s+t,R} a_r b_s c_{t,n}, \tag{56}$$

with coefficients defined by the expansions

$$y_0^{-\varepsilon} = \sum_{r=0}^{\infty} a_r \varepsilon^r,$$

$$\frac{\pi \varepsilon}{\sin \pi \varepsilon} \sim \sum_{s=0}^{\infty} b_s \varepsilon^s,$$

$$\frac{n!}{\Gamma(n+1-\varepsilon)} \sim \sum_{t=0}^{\infty} c_{t,n} \varepsilon^s. \tag{57}$$

Expressions for a_r, b_s are readily found,

$$a_r = \frac{(-\ln y_0)^r}{r!},$$

$$b_s = \begin{cases} \frac{2(2^{s-1}-1)\pi^s |B_s|}{s!}, & s \text{ even}, \\ 0, & s \text{ odd}, \end{cases} \tag{58}$$

where B_s are Bernoulli numbers. As for $c_{t,n}$, it can be written as a sum over Young tableaux of order t,

$$c_{t,n} = \sum_{Y.T.(t)} (-1)^{t - \sum_{j=1}^{t} v_j} \frac{\prod_{i=1}^{t} [\psi^{(i-1)}(n+1)]^{v_i}}{\prod_{\ell=1}^{t} \ell!^{v_\ell} v_\ell!}, \tag{59}$$

where v_j, $j = 1, \ldots, t$, are the numbers of rows with j elements, such that $\sum_{j=1}^{t} j v_j = t$, and ψ is the digamma function.

As an example, for the $k = 1$ sector in the $\mathbb{CP}^1$ model we need

$$T_{1,n} = \psi(n+1) - \ln y_0, \tag{60}$$

from which Eq. (54) obtains. For the $k = 2$ sector in the same model, expression (59) is already too cumbersome to compute (56) explicitly:

$$T_{3,n} = -\frac{1}{6}(\ln y_0)^3 - \frac{\pi^2}{6}\ln y_0 + \left[\frac{1}{2}(\ln y_0)^2 + \frac{\pi^2}{6}\right]\psi(n+1)$$

$$- (\ln y_0)\left(\frac{1}{2}\psi(n+1)^2 - \psi'(n+1)\right)$$

$$+ \frac{1}{6}\psi(n+1)^3 - \frac{1}{2}\psi(n+1)\psi'(n+1) + \frac{1}{6}\psi''(n+1). \tag{61}$$

However, it is not necessary to perform the summation in (55) to realise that the pole structure has a natural interpretation in terms of instantons. From the original classical action

$$S[\Psi, \Psi^\dagger, A_\mu, \alpha] = \frac{N}{2g_0^2} \int_{T^2} d^2x \, |D_\mu \Psi|^2 + \frac{N}{2g_0^2} \int_{T^2} d^2x \, \alpha(x)(\Psi^\dagger \Psi - 1), \tag{62}$$

the classical equations of motion

$$(-D_\mu^2 + \alpha)\Psi = 0, \quad \alpha = \Psi^\dagger D_\mu^2 \Psi \tag{63}$$

ensure that, for classical solutions, the value of the action is given by the integral of $-\alpha$:

$$S_{cl} = \frac{N}{2g_0^2} \int_{T^2} d^2x \, |D_\mu \Psi|^2 = \frac{N}{2g_0^2} \int_{T^2} d^2x \, (-\alpha). \tag{64}$$

The nth pole of the integrand of Z_k corresponds to a value of $\alpha = m^2$ such that

$$\frac{m^2 L^2}{4\pi |k|} = -\left(n + \frac{1}{2}\right) \quad \Longrightarrow \quad S_{\text{cl}} = \frac{N}{2g_0^2} m^2 L^2 = (1 + 2n)\frac{N}{2g_0^2} 2\pi |k|, \tag{65}$$

i.e., it exactly matches the classical action of a *multiinstanton* configuration composed of an instanton and n instanton–antiinstanton pairs (Fig. 4). This is in contrast with the case of the sphere, where the structure of poles of the integrand of Z_k does not correspond to charge k multiinstanton configurations.

Notice that, although there are no unit charge instantons on the torus [34], there is a nontrivial contribution of the $k = \pm 1$ sectors to the total partition function. This is so because the dominant configurations in the partition function are not the smooth classical solutions, but rather the rough configurations in the neighbourhood of these. The classical solution fails to exist for unit charge, but this is a classical accident, which does not affect the quantum dynamics of the system because of the nonzero contribution of the neighbouring configurations.

4.2.2. The saddle point method

Now let us consider the extrema of $\widetilde{S}(y)$. These are the zeros of its derivative,

$$\frac{d\widetilde{S}(y)}{dy} = \ln y_0 - \psi(y + 1/2), \tag{66}$$

and constitute a sequence $y = s_n$, $n = 0, 1, \ldots$. Saddle points and poles alternate, $s_0 > p_0 > s_1 > p_1 > \cdots$, as seen in Figs. 3 and 4. The saddle points approach the poles for large n, $\lim_{n \to \infty} s_n / p_n = 1$.

If the dominant saddle point s_0 lies in the region $y \gg 1$, we can find its location as an expansion in powers of y_0^{-1} starting from (47):

$$s_0 = y_0 + \frac{1}{24 y_0} - \frac{12097}{576 y_0^3} + \mathcal{O}(y_0^{-5}). \tag{67}$$

Equivalently, the infinite volume value m_{T}^2 of the saddle point receives finite volume corrections,

$$m_{s_0}^2 = m_{\text{T}}^2 \left(1 + \frac{2\pi^2 |k|}{3 m_{\text{T}}^4 L^4} + \mathcal{O}(y_0^{-6})\right). \tag{68}$$

We evaluate the partition function in sector k, up to quadratic order,

$$Z_k^{(s_0)} \approx \frac{4\pi |k|}{L^2} \sqrt{\frac{2\pi}{N|k|}} e^{-N|k|\widetilde{S}(s_0)} \left(\frac{d^2 \widetilde{S}}{dy^2}\right)_{s_0}^{-1/2}, \tag{69}$$

from which

$$\begin{aligned}
Z_k^{(s_0)} &= \frac{4\pi |k|}{L^2} \sqrt{\frac{2\pi y_0}{N|k|}} \exp\left\{-N|k| y_0 - \frac{N|k|}{24 y_0} + \frac{1}{16 y_0^2} + \frac{29 N|k|}{5760 y_0^3} + \mathcal{O}(y_0^{-5})\right\} \\
&= \frac{4\pi m_{\text{T}}}{\sqrt{2N} L} \exp\left\{-N\frac{m_{\text{T}}^2 L^2}{4\pi} - \frac{\pi}{6}\frac{N k^2}{m_{\text{T}}^2 L^2} + \left(\frac{\pi |k|}{m_{\text{T}}^2 L^2}\right)^2 + \mathcal{O}(y_0^{-3})\right\}.
\end{aligned} \tag{70}$$

The partition function in the trivial topological sector can be computed in the large $m^2 L^2$ limit by substituting an integral for the sum in the definition of the ζ function,

$$Z_0 = \int dm^2 \exp\left\{N\frac{m^2 L^2}{4\pi}\left(\ln \frac{m^2}{m_{\text{T}}^2} - 1\right)\right\}, \tag{71}$$

and coincides with the $B \to 0$ limit of the result for Z_k. There are no poles in this sector, which is compatible with the absence of instantons of zero topological charge, and its only saddle point yields an approximation

$$Z_0^{(s_0)} \approx \frac{4\pi m_{\mathrm{T}}}{\sqrt{2N}L} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\}, \tag{72}$$

agreeing with the result (70) for the nontrivial sectors.

The full partition function in terms of the vacuum angle θ, in the saddle point approximation, is the sum

$$
\begin{aligned}
Z^{(s_0)}(\theta) &= \sum_{k\in\mathbb{Z}} Z_k^{(s_0)} e^{-ik\theta} \\
&\approx \frac{4\pi m_{\mathrm{T}}}{\sqrt{2N}L} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\} \sum_{k\in\mathbb{Z}} \exp\left\{-\frac{\pi}{6}\frac{Nk^2}{m_{\mathrm{T}}^2 L^2} - ik\theta\right\} \\
&= \frac{4\pi m_{\mathrm{T}}}{\sqrt{2N}L} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\} \vartheta_3\left(\frac{\theta}{2\pi}\,\middle|\,\frac{iN}{6m_{\mathrm{T}}^2 L^2}\right).
\end{aligned}
\tag{73}
$$

Still in the region $y \gg 1$, i.e. $m^2 L^2 \gg 4\pi|k|$, let us analyse this partition function in two different regimes, depending on the relation between the number N and $m_{\mathrm{T}}^2 L^2$.

For $N \gg m_{\mathrm{T}}^2 L^2$, the contribution of high topological sectors (large $|k|$) to (73) can be neglected. Keeping just sectors $k = 0, \pm 1$, we obtain

$$Z^{(s_0)}(\theta) \approx \frac{4\pi m_{\mathrm{T}}}{\sqrt{2N}L} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\} \left(1 + 2\exp\left\{-\frac{\pi}{6}\frac{N}{m_{\mathrm{T}}^2 L^2}\right\}\cos\theta\right), \tag{74}$$

giving rise to a vacuum energy density

$$
\begin{aligned}
\mathcal{E}_0^{(s_0)}(\theta) - \mathcal{E}_0^{(s_0)}(0) &= -\frac{1}{L^2}\ln\frac{Z^{(s_0)}(\theta)}{Z^{(s_0)}(0)} \\
&\approx \frac{4}{L^2}\exp\left\{-\frac{\pi}{6}\frac{N}{m_{\mathrm{T}}^2 L^2}\right\}(1 - \cos\theta),
\end{aligned}
\tag{75}
$$

i.e., the typical $\mathcal{E}_0(\theta)$ dependence of a dilute instanton gas. This is a 2π-periodic function, smooth for all values of θ including $\theta = \pm\pi$. Hence, there is no first order phase transition.

This regime would be compatible with a *definition* of the $1/N$ expansion in which both the large N limit and the $1/N$ corrections are studied in finite volume.

However, if we move to the region $m_{\mathrm{T}}^2 L^2 \gg N$, it proves convenient to use the Poisson formula in (73) to get

$$
\begin{aligned}
Z^{(s_0)}(\theta) &= \frac{4\sqrt{3}\pi m_{\mathrm{T}}^2}{N} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\} \vartheta\left[\begin{array}{c}\frac{\theta}{2\pi}\\0\end{array}\right]\left(0\,\middle|\,\frac{i6m_{\mathrm{T}}^2 L^2}{N}\right) \\
&= \frac{4\sqrt{3}\pi m_{\mathrm{T}}^2}{N} \exp\left\{-N\frac{m_{\mathrm{T}}^2 L^2}{4\pi}\right\} \sum_{q\in\mathbb{Z}} \exp\left\{-\frac{6\pi m_{\mathrm{T}}^2 L^2}{N}\left(q + \frac{\theta}{2\pi}\right)^2\right\},
\end{aligned}
\tag{76}
$$

where the ϑ function with characteristics is given by

$$\vartheta\left[\begin{array}{c}a\\b\end{array}\right](z|\tau) = \sum_{n\in\mathbb{Z}} \exp\left\{i\pi\tau(n+a)^2 + i2\pi(n+a)(z+b)\right\}. \tag{77}$$

Now, it is the dual sum in q that is dominated by the low $|q|$ terms. Defining again $\widetilde{\theta}$ as the angle in $[-\pi, +\pi]$ differing from θ by an integer,

$$Z^{(s_0)}(\theta) \approx \frac{4\sqrt{3}\pi m_T^2}{N} \exp\left\{-N\frac{m_T^2 L^2}{4\pi} - \frac{3m_T^2 L^2}{2\pi N}\widetilde{\theta}^2\right\}. \tag{78}$$

This reproduces the traditional large N picture, where instanton effects are suppressed, and the vacuum energy density depends quadratically on θ within the interval $[-\pi, +\pi]$:

$$\mathcal{E}_0^{(s_0)}(\theta) - \mathcal{E}_0^{(s_0)}(0) \approx \frac{3m_T^2}{2\pi N}\widetilde{\theta}^2. \tag{79}$$

Periodicity in θ is guaranteed because $\mathcal{E}_0(\theta)$ depends on the periodic variable $\widetilde{\theta}$, but this function is not smooth at odd multiples of π, where $\widetilde{\theta}$ is doubly defined, levels cross and a first order phase transition occurs.

This regime is compatible with a definition of the $1/N$ expansion in which the thermodynamic limit is performed first, and N is taken to infinity afterwards.

These results agree with the analysis of $\mathbb{C}P^{N-1}$ models on the sphere. Since the physical pictures pertaining to the regimes $N \gg m_T^2 L^2$ and $m_T^2 L^2 \gg N$ are different, the limits $N \to \infty$ and $V \to \infty$ do not commute, and the orders in which these limits are taken determine different theories. This behaviour points towards the existence of a phase transition in the $N \to \infty$ theory.

5. Discussion

After the careful analysis of the large N method on the sphere and the torus, we conclude that the apparent incompatibility between instanton physics and the $1/N$ expansion has its cause in the formulation of the latter in infinite volume and is a subtle effect of the noncommutativity of the large N and thermodynamic limits.

To clarify this, consider the essential dependence of the vacuum energy density on the angle θ, the volume V, and the number of colours N, at fixed saddle point mass m_T

$$\mathcal{E}_0(\theta) = -\frac{1}{V}\ln\frac{\vartheta_3\left(\frac{\theta}{2\pi}\Big|\frac{iN}{6m_T^2 V}\right)}{\vartheta_3\left(0\Big|\frac{iN}{6m_T^2 V}\right)}, \tag{80}$$

which is valid in the cases of the sphere and the torus provided we *define* the free energy as a function of θ by subtraction of the contribution at $\theta = 0$ for each N.

The various limits of (80) are best discussed in terms of a dimensionless variable $x \equiv N/(6m_T^2 V)$ and the function

$$\frac{N}{6m_T^2}\mathcal{E}_0(\theta) \equiv f(\theta, x) = -x\ln\frac{\vartheta_3\left(\frac{\theta}{2\pi}\big|ix\right)}{\vartheta_3(0|ix)} = -x\ln\frac{\vartheta\left[\begin{smallmatrix}\theta/(2\pi)\\0\end{smallmatrix}\right]\left(0\big|\frac{i}{x}\right)}{\vartheta_3\left(0\big|\frac{i}{x}\right)}. \tag{81}$$

The last equation is obtained by applying the Poisson resummation formula, that is, the modular transformation of the theta functions. For x strictly positive, f is a well defined real analytic function of $\theta \in \mathbb{R}$. The nonanalyticities for complex θ are branch cuts located at the zeros of the Jacobi ϑ_3 function, that is, for $\theta = (1 + ix)\pi$ and its translations by integer multiples of 2π and of $i2\pi x$. These zeros never occur for real values of θ.

The two limiting regimes we have been discussing are given in terms of the dimensionless quantities as:

- Semiclassical: send x to infinity (meaning $N \gg m_T^2 V$). In this case

$$\vartheta_3\left(\frac{\theta}{2\pi}\bigg| ix\right) = 1 + 2e^{-\pi x}\cos\theta + \mathcal{O}(e^{-2\pi x}), \tag{82}$$

and we recover the nonperturbative result

$$f(\theta, x) = xe^{-\pi x}(1 - \cos\theta) + \mathcal{O}(e^{-2\pi x}), \tag{83}$$

equivalent to

$$\mathcal{E}_0(\theta) = \frac{1}{V}\exp\left\{-\frac{\pi N}{6m_T^2 V}\right\}(1 - \cos\theta) + \mathcal{O}\left(\frac{m_T^2}{N}\exp\left\{-\frac{\pi N}{3m_T^2 V}\right\}\right), \tag{84}$$

equivalent to (2) (and vanishing as $x \to \infty$ together with all of its derivatives).
- Traditional large N: send x to zero (meaning $N \ll m_T^2 V$). This is a problematic limit for the modular parameter $\tau = ix$, which leaves the upper half plane. We consider separately the regions $\theta \in (-\pi, \pi)$ and $\theta = \pi$ (the rest of the function is obtained by periodicity):

$$f(\theta \in (-\pi, \pi), x) = \frac{\theta^2}{4\pi} + \mathcal{O}\left(\exp\left\{-\frac{\pi(1 - \theta/\pi)^2}{x}\right\}\right),$$

$$f(\theta = \pi, x) = \frac{\pi^2}{4\pi} - x\ln 2 + \mathcal{O}(e^{-\pi/x}), \tag{85}$$

reproducing (1), including the first order cusp at $\theta = \pi$:

$$\mathcal{E}_0(\theta \in (-\pi, \pi)) = \frac{3m_T^2}{2\pi N}\theta^2 + \mathcal{O}\left(\frac{m_T^2}{N}\exp\left\{-\frac{6\pi m_T^2 V(1 - \theta/\pi)^2}{N}\right\}\right),$$

$$\mathcal{E}_0(\theta = \pi) = \frac{3m_T^2}{2\pi N}\pi^2 - \frac{\ln 2}{V} + \mathcal{O}\left(\frac{m_T^2}{N}\exp\left\{-\frac{6\pi m_T^2 V}{N}\right\}\right). \tag{86}$$

The rôle of ix as a modular parameter suggests an analogy with finite temperature models, where the number of colours corresponds to the inverse temperature β. In this sense, we expect that the physics at $1/N = 0$ corresponds to zero temperature phenomena. The fact that the thermodynamic and large N limits do not commute (reflected in the behaviour of x) would suggest the presence of a phase transition exactly at zero temperature, i.e., $1/N = 0$.

Returning to physical quantities, we have shown that finite volume effects in the θ dependence of the vacuum energy density for $\mathbb{CP}^{N-1}$ models on S^2 and T^2, when all topological sectors are taken into account, give rise to two asymptotic regimes, one dominated by instanton effects (when $N \gg m_T^2 V$) and the other by the conventional large N picture (when $N \ll m_T^2 V$). These are smoothly connected by an interpolating region.

It should be realised that the basic hypotheses of the method of large N do not hold when $N \ll m_T^2 V$, for which precisely the traditional large N results obtain. The saddle point technique needs N to be the largest dimensionless parameter of the theory, in particular larger than $m_T^2 V$. Two very different theories are defined by interchanging the noncommuting limits $N \to \infty$ and $V \to \infty$. In principle, the only procedure consistent with the saddle point method is taking the large N limit first, and then going over to the thermodynamic limit.

However, for small values of θ, where the large N approximation is expected to hold [41], both procedures seem to make sense. Lattice measurements of the $\theta = 0$ topological susceptibility agree with the traditional large N picture, corresponding to performing first the $V \to \infty$ limit,

Table 1
Numerical values of the mass gap m for different $\mathbb{CP}^{N-1}$ models at different volumes $m^2 L^2$. The models are in the thermodynamic regime $m^2 L^2 > N$ in all cases. The rôle of instantons is only manifest in the cases $\mathbb{CP}^1$ and $\mathbb{CP}^2$ [22].

Reference	m	L	$m^2 L^2$	N
[42] Blatter et al.	0.165	40	43.6	2
[43] Ahmad et al.	0.179	50	80.1	2
[22] Lian–Thacker	0.111	100	123.2	2
[44] Keith–Hynes–Thacker	0.179	50	80.1	2
[22] Lian–Thacker	0.084	100	70.6	3
[45] Campostrini et al.	0.066	120	62.7	4
[47] Burkhalter et al.	0.131	32	17.6	4
[22] Lian–Thacker	0.088	100	77.4	4
[44] Keith–Hynes–Thacker	0.180	50	81.0	4
[43] Ahmad et al.	0.186	50	86.5	6
[22] Lian–Thacker	0.085	100	72.2	6
[45] Campostrini et al.	0.196	72	199.1	10
[48] Del Debbio et al.	0.192	60	132.7	10
[43] Ahmad et al.	0.212	50	112.4	10
[22] Lian–Thacker	0.058	100	33.6	10
[44] Keith–Hynes–Thacker	0.212	50	112.4	10
[48] Del Debbio et al.	0.397	42	278.0	15
[48] Del Debbio et al.	0.418	30	157.3	21
[49] Vicari	0.287	60	296.5	21
[48] Del Debbio et al.	0.305	56	291.8	21
[49] Vicari	0.411	42	298.0	41

and then taking N to infinity. However, these simulations have been only carried out for values of the parameters such that $m_{\mathrm{T}}^2 V \gtrsim N$ (see Table 1 and Refs. [46,50,51]), agreeing with our analysis in the region not validated by the saddle point method. It would be interesting to rerun these simulations for smaller volumes, still close to the thermodynamic limit with a stable mass gap, but where an instanton-dominated θ dependence of the vacuum energy density could emerge and eventually take over. Notice that the singularity of the topological susceptibility pointed out by lattice simulations in the $\mathbb{CP}^1$ model is also found in the semiclassical scenario. The existence of such a singularity can also be understood by the presence of a family of Lee–Yang zeros of the analytic continuation of the partition function in the complex θ-plane, converging to $\theta = 0$ in the thermodynamic limit [9].

As regards the neighbourhood of $\theta = \pi$, the nodal analysis of [13] appears incompatible with a first order phase transition at $\theta = \pi$. This is the behaviour of the system for lower values of N for any volume, i.e. $\mathbb{CP}^1$ and $\mathbb{CP}^2$ models. For larger values of N, for instance $N > 4$, this behaviour is only observed for small volumes, i.e. volumes which verify $m^2 V \ll N$. For larger volumes the effect is swept off by the infrared fluctuations and the system undergoes a phase transition at $\theta = \pi$ with spontaneous CP symmetry breaking.

The behaviour for intermediate values of θ has so far proved elusive to numerical techniques, due to the inaccuracies inherent to lattice simulations in this region. Let us however remark that a novel technique [52] based on analytically continuing the θ dependence to imaginary values has been introduced to overcome this problem. This technique has been applied to the $\mathbb{CP}^9$ model [53], with conclusions agreeing with the usual large N expansion. The consistency of this technique for any value of N is supported by the absence of singularities in the analytic extension of the partition function to the whole θ-plane [9].

Summing up, the large N method is compatible with instanton effects. Besides, the results on θ dependence obtained with this tool agree with the known behaviour at $\theta = 0$, i.e., the vacuum energy density is differentiable there and the Vafa–Witten theorem holds. It is also compatible with the numerical determination of the topological susceptibility at $\theta = 0$. The analysis of the poles of the partition function on the torus supports the method of Jevicki, whereby instantonic effects appear in the large N limit in the form of resonances. First order phase transitions with spontaneous parity breaking at $\theta = \pi$ appear in the formulation of the models directly at infinite volume, and we have exposed the analytic roots of this fact. The large N method is thus in agreement with all exact results on θ dependence, and provides a valuable bridge between the angles $\theta = 0$ and $\theta = \pi$.

Let us remark that the behaviour of the theory in finite volume plays a fundamental rôle in condensed matter settings, where sigma models can be used as effective theories for the quantum Hall effect [54–57]. In this context, the Hall conductivity is identified with the coupling of the topological term, and the stability of Hall plateaux is linked to the renormalisation group running of the couplings (including θ). The large N limit of $\mathbb{CP}^{N-1}$ models was studied in connection with this phenomenon in a series of papers (see, e.g., [58–62]), in which the different regimes we have discussed were also identified; in this case the traditional large N limit at infinite volume is blind to edge effects, which of course are crucial for the physics of the Hall effect. In particular, edge currents are a finite size effect and this suggests that the $m^2 V \ll N$ regime of $\mathbb{CP}^{N-1}$ sigma models is the relevant regime for their description.

Finally, we remark that the difference between the two regimes is due to the asymptotic behaviour of the effective action in the different topological sectors. According to Eqs. (48) and (50), for small values of the topological charge, $|k| < m^2 V$, the effective action is quadratic in k, whereas for large topological charges its leading term is linear in $|k|$ [9]. The two regimes also differ at finite temperature. Since the spacetime volume is $V = LT$, the change of asymptotic dependence of the effective action on the topological charge can be associated with a finite temperature crossover from the low temperature regime $\beta = 1/T > m^2 L/|q|$ to the high temperature regime $\beta < m^2 L/|q|$ and cannot be related to any phase transition [63]. One might expect a similar phenomenon in QCD, although in that case there is a finite temperature phase transition [64].

Acknowledgements

References

[1] H. Eichenherr, Nucl. Phys. B 146 (1978) 215.
[2] V.L. Golo, A.M. Perelomov, Phys. Lett. B 79 (1978) 112.
[3] A. d'Adda, M. Lüscher, P. Di Vecchia, Nucl. Phys. B 146 (1978) 63.
[4] E. Vicari, H. Panagopoulos, Phys. Rep. 470 (2009) 93.
[5] M. Asorey, J.G. Esteve, A.F. Pacheco, Phys. Rev. D 27 (1983) 1852.
[6] P.B. Wiegmann, Phys. Lett. B 152 (1985) 209.
[7] F.D.M. Haldane, Phys. Rev. Lett. 50 (1983) 1153.
[8] C. Vafa, E. Witten, Phys. Rev. Lett. 53 (1984) 535.

  [9] M. Aguado, M. Asorey, Phys. Rev. D 80 (2009) 127702.
 [10] A.B. Zamolodchikov, A.B. Zamolodchikov, Nucl. Phys. B 379 (1992) 602.
 [11] R. Shankar, N. Read, Nucl. Phys. B 336 (1990) 457.
 [12] E. Lieb, T. Schultz, D. Mattis, Ann. Phys. 16 (1961) 407.
 [13] M. Asorey, F. Falceto, Phys. Rev. Lett. 80 (1998) 234.
 [14] M. Asorey, in: Proceedings of the Conference on Geometry and Physics, CIEMAT, Madrid, 1998.
 [15] M. Asorey, F. Falceto, Phys. Rev. Lett. 77 (1996) 3074.
 [16] G. 't Hooft, Phys. Rev. D 14 (1976) 3432.
 [17] D. Diakonov, V.Y. Petrov, Nucl. Phys. B 245 (1984) 259.
 [18] D. Diakonov, V.Y. Petrov, Nucl. Phys. B 272 (1986) 457.
 [19] E.V. Shuryak, Nucl. Phys. B 198 (1982) 83.
 [20] M. Lüscher, Nucl. Phys. B 200 (1982) 61.
 [21] B. Berg, M. Lüscher, Nucl. Phys. B 190 (1981) 412.
 [22] Y. Lian, H.B. Thacker, Phys. Rev. D 75 (2007) 065031.
 [23] G. 't Hooft, Nucl. Phys. B 75 (1974) 461.
 [24] G. 't Hooft, Nucl. Phys. B 72 (1974) 461.
 [25] E. Witten, Nucl. Phys. B 149 (1979) 285.
 [26] A. Jevicki, Phys. Rev. D 20 (1979) 3331.
 [27] J.M. Maldacena, Adv. Theor. Math. Phys. 2 (1998) 231.
 [28] E. Witten, Nucl. Phys. B 156 (1979) 269.
 [29] G. Veneziano, Nucl. Phys. B 159 (1979) 213.
 [30] P. Schwab, Phys. Lett. B 118 (1982) 373.
 [31] P. Schwab, Phys. Lett. B 126 (1983) 241.
 [32] G. Münster, Phys. Lett. B 118 (1982) 380.
 [33] G. Münster, Nucl. Phys. B 218 (1983) 1.
 [34] P.J. Braam, P. van Baal, Commun. Math. Phys. 122 (1989) 267.
 [35] M. Aguado, Ph.D. thesis, U. Zaragoza, 2003.
 [36] M. Aguado, M. Asorey, D. García-Álvarez, Mod. Phys. Lett. A 18 (2003) 2303.
 [37] M. Asorey, F. Falceto, G. Luzón, Contemp. Math. 219 (1998) 1.
 [38] M. Aguado, M. Asorey, J.G. Esteve, Commun. Math. Phys. 218 (2001) 233.
 [39] S.K. Blau, M. Visser, A. Wipf, Int. J. Mod. Phys. A 6 (1991) 5409.
 [40] M. Aguado, M. Asorey, A. Wipf, Annals Phys. 298 (2002) 2.
 [41] S.R. Coleman, Ann. Phys. 101 (1976) 239.
 [42] M. Blatter, R. Burkhalter, P. Hasenfratz, F. Niedermayer, Phys. Rev. D 53 (1996) 923.
 [43] S. Ahmad, J.T. Lenaghan, H.B. Thacker, Phys. Rev. D 72 (2005) 114511.
 [44] P. Keith-Hynes, H.B. Thacker, Phys. Rev. D 78 (2008) 025009.
 [45] M. Campostrini, P. Rossi, E. Vicari, Phys. Rev. D 46 (1992) 4643.
 [46] R. Burkhalter, Fixed point action and topology in the CP3 model, Phys. Rev. D 54 4121 (1996).
 [47] R. Burkhalter, M. Imachi, Y. Shinno, H. Yoneyama, Progr. Theoret. Phys. 106 (2001) 613.
 [48] L. Del Debbio, G. Manca, E. Vicari, Phys. Lett. B 594 (2004) 315.
 [49] E. Vicari, Phys. Lett. B 309 (1993) 139.
 [50] B. Alles, L. Cosmai, M. D'Elia, A. Papa, Nucl. Phys. Proc. Suppl. 83 (2000) 530.
 [51] M. Campostrini, P. Rossi, E. Vicari, Nucl. Phys. Proc. Suppl. 30 (1993) 830.
 [52] V. Azcoiti, G. Di Carlo, A. Galante, V. Laliena, Phys. Rev. Lett. 89 (2002) 141601.
 [53] V. Azcoiti, G. Di Carlo, A. Galante, V. Laliena, Phys. Rev. D 69 (2004) 056006.
 [54] H. Levine, S.B. Libby, A.M.M. Pruisken, Phys. Rev. Lett. 51 (1983) 1915.
 [55] A.M.M. Pruisken, Nucl. Phys. B 235 (1984) 277.
 [56] D.E. Khmel'nitskiǐ, JETP Lett. 38 (1983) 552.
 [57] H. Levine, S.B. Libby, A.M.M. Pruisken, Nucl. Phys. B 240 (1984) 30;
      H. Levine, S.B. Libby, A.M.M. Pruisken, Nucl. Phys. B 240 (1984) 49;
      H. Levine, S.B. Libby, A.M.M. Pruisken, Nucl. Phys. B 240 (1984) 71.
 [58] A.M.M. Pruisken, M.A. Baranov, M. Voropaev, Phys. Rev. Lett. 505 (2003) 4432.
 [59] A.M.M. Pruisken, I.S. Burmistrov, Annals Phys. (N.Y.) 316 (2005) 285;
      A.M.M. Pruisken, I.S. Burmistrov, Annals Phys. (N.Y.) 322 (2007) 1265.
 [60] A.M.M. Pruisken, R. Shankar, N. Surendran, Phys. Rev. B 72 (2005) 035329.
 [61] I.S. Burmistrov, A.M.M. Pruisken, Phys. Rev. Lett. 101 (2008) 056801.

[62] A.M.M. Pruisken, Int. J. Theor. Phys. 48 (2009) 1736.
[63] I. Affleck, Nucl. Phys. B 162 (1980) 461;
     I. Affleck, Nucl. Phys. B 171 (1980) 420.
[64] D. Kharzeev, R.D. Pisarski, M.H.G. Tytgat, Phys. Rev. Lett. 81 (1998) 512.