Nuclear Physics B144 (1978) 513–524

NATURE OF THE PHASE TRANSITION IN A NON-LINEAR $O(2)_3$ MODEL

Paul R. THOMAS and Michael STONE
DAMTP, Silver street, Cambridge, UK

Received 5 May 1978

We examine the gas of topological excitations in the lattice non-linear $O(2)_3$ model. We argue that the phase transition between broken and unbroken symmetries can be identified with that of three-dimensional scalar electrodynamics, strongly suggesting that the transition is first order.

1. Introduction

In a recent paper by Banks, Kogut and Myerson [1], the problem is posed of calculating the critical exponents of the three-dimensional O(2) Heisenberg model by studying a gas of its topological excitations. These topological excitations are vortex loops, the three-dimensional analogue of the Kosterlitz and Thouless vortices [2]. This paper is a report on our attempts to do this. We have failed because the phase transition appears to be first order and therefore not dominated by a critical point.

To be more precise, we investigated an $O(2)_3$ Villain model, generalised in the same spirit as Jose et al. [3]. We find that this model is dual to a three-dimensional model of scalar electrodynamics (or the Landau-Ginzberg model of superconductivity), which is known to have a first-order phase transition between the normal and Higgs-Kibble (superconducting) phases [4]. This duality is interesting because it is a topological excitation to particle duality of the type proposed by Olive [5] *. The physical reason for the first-order phase transition is not hard to find, being analogous to the cooperative interaction between dislocations in solids that makes melting a first-order transition [6] (a better known example than superconductivity!) In fact this model is probably a reasonable one for some types of dislocation melting.

The $O(2)_3$ Heisenberg model is an extremum of the family of generalised $O(2)_3$ models and the duality transformation in the Villain model is singular. However, a continuum version of the duality map [7] seems to suggest that scalar QED is the right model to discuss, so we are unsure about the nature of the transition in this case.

The layout of the paper is as follows. In sect. 2 we will investigate the Villain

* Also the sine-Gordon-Thirring equivalence.

models and in particular illustrate the duality mapping. In sect. 3 we introduce a "continuum" formalism for vortex-loop gases. In sect. 4 we will argue that the transition is first order and discuss the physical reason for this. In the latter section we will give a general discussion of why these $O(2)_3$ models are not in the obvious universality class of the n vector model.

2. Exact duality maps

All the models discussed here are of the Villain type [8] which have the nice property that the partition function factorises into a spin-wave and a topological-excitation part

$$Z = Z_{\text{spin wave}} Z_{\text{topological}} \ . \tag{2.1}$$

There is quite extensive literature developing these models [1,3,9] and the reader is referred to these papers for further details and examples.

The Villain $O(2)_3$ model is defined by

$$Z[J] = \int_{-\pi}^{+\pi} \mathrm{d}\left[\frac{\theta_i}{2\pi}\right] \sum_{\{n_\mu\}} \exp[-\tfrac{1}{2}\beta \sum \{(\Delta_\mu \theta(i) + 2\pi n_\mu)^2 + iJ(i)\,\theta(i)\}] \ , \tag{2.2}$$

where

$$\Delta_\mu \theta(i) = \theta(i+\hat{\mu}) - \theta(i) \ .$$

The variables n_μ are defined on the directed links between adjacent sites and the J's are integer valued external sources. The spin correlation function may be easily shown to be

$$\langle \cos(\theta(R) - \theta(0)) \rangle = Z[J]/Z[0] \ , \tag{2.3}$$

with

$$J = \delta(x-R) - \delta(x) \ .$$

Using the Poisson resummation formula, we may easily rewrite (2.2) as

$$Z = \int_{-\pi}^{+\pi} \mathrm{d}\left[\frac{\theta(i)}{2\pi}\right] \frac{1}{[2\pi\beta]^{N/2}} \sum_{\{n_\mu\}} \exp\left[-\frac{1}{2\beta} \sum_{i,\mu} \{n_\mu^2 + i\Delta_\mu \theta(i) + iJ(i)\,\theta(i)\}\right] , \tag{2.4}$$

where N is the number of lattice sites. Doing the θ integration gives:

$$Z[J] = (2\pi\beta)^{-N/2} \sum_{\{n_\mu\}} \delta_{\Delta_\mu n_\mu,\,J} \exp(-\sum_i n_\mu^2/2\beta) \ . \tag{2.5}$$

The factor $(2\pi\beta)^{-N/2}$ is the spin-wave part in the absence of sources. The condition that $\Delta_\mu n_\mu = J$ implies that the partition function is a sum over threads of action

$1/2\beta$ per unit length which either form closed loops or terminate on the sources J. We will later interpret these threads as the Nielson-Oleson [10] (or Abrikosov) vortex lines of scalar QED or the particle trajectories of the field theory associated with the $O(2)_3$ model. With this interpretation it is clear that the J's act as monopole sources for the vortex lines.

We proceed exactly as in ref. [1] by introducing integer valued variables m_μ, φ_μ, living on the links of the dual lattice, such that

$$n_\mu = \epsilon_{\mu\nu\alpha}\Delta_\gamma m_\alpha + \varphi_\mu .$$
$$\Delta_\mu \varphi_\mu = J . \tag{2.6}$$

The m_μ are arbitrary up to the addition of lattice gradients, so one has to choose a gauge or divide out by the group volume of $\mathbb{R}$. We will see later that the choice of φ_μ is also arbitrary but for the moment visualise them as flux lines joining up the monopole-antimonopole pairs.

We replace the sum over m_μ by integration over a continuum vector field A_μ by further use of the Poisson formula to obtain

$$Z = (2\pi\beta)^{-N/2}\Omega^{-1}\int \mathrm{d}[A_\mu] \sum_{\{l_\mu\}} \delta_{\Delta_\mu l_\mu, 0}$$
$$\times \exp \sum_i \{ -(\Delta \wedge A + \varphi)^2/2\beta + 2\pi i l_\mu A_\mu \} . \tag{2.7}$$

l_μ also lies on the links of the dual lattice, Ω is the group volume and the constraint on $\Delta_\mu l_\mu$ arises because the volume factor suppresses terms with $\Delta_\mu l_\mu \neq 0$ (this constraint, i.e. current conservation for the sources of A_μ, would also follow from any choice of gauge; clearly assuming it makes any choice of gauge irrelevant).

If we make some different choice of φ_μ satisfying $\Delta_\mu \varphi_\mu = J$, then the difference

$$\varphi_\mu^{(1)} - \varphi_\mu^{(2)} = \Delta \wedge A', \quad A' \in \mathbb{Z} , \tag{2.8}$$

and replacing A_μ by $A_\mu + A'_\mu$ would leave the partition function invariant because

$$\mathrm{e}^{2\pi i A'_\mu} = 1 . \tag{2.9}$$

This is a statement of the Dirac condition relating the magnitude of the monopole charge, J, and the magnitude of the current loops.

One can easily see that adding a flux tube (Dirac string) to the field $\Delta \wedge A$ is the correct way to insert monopole sources into a field theory described by divergenceless fields. If one considers a monopole-antimonopole pair then expanding out the gauge action we find

$$\tfrac{1}{2}(\Delta \wedge A + \varphi)^2 = \tfrac{1}{2}(\Delta \wedge A)^2 + \tfrac{1}{2}\varphi^2 + (\Delta \wedge A) \cdot \varphi$$
$$= \tfrac{1}{2}(\Delta \wedge A)^2 + \tfrac{1}{2}\varphi^2 + (\Delta \wedge \varphi) \cdot A$$
$$= \tfrac{1}{2}(\Delta \wedge A)^2 + \tfrac{1}{2}\varphi^2 + A \cdot J , \tag{2.10}$$

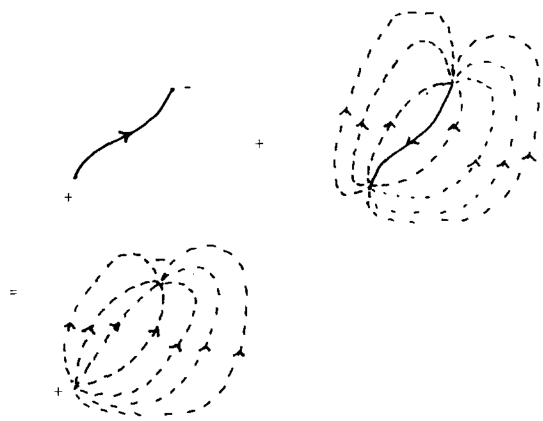

Fig. 1. The string field φ s cancelled by the solenoidal field A to give the correct dipole field.

where J is a solenoidal source creating an A field cancelling off the string and substituting the correct dipole field (see fig. 1).

Returning to eq. (2.7) we observe that the A_μ integration yields long range forces of the Biot-Savart type between the loops described by l_μ. These loops are readily interpreted as the vortex threads of the original $O(2)_3$ model.

The alert reader might object at this point because in order to obtain (2.7) we have integrated over all configurations in the $O(2)_3$ model. So how can we identify vortices each distribution of which is a function of the initial configuration of $\theta(l)$'s? In fact it is possible to proceed directly from (2.2) to (2.8) without summing over configurations or using the Poisson summation formula. For details of this see appendix D of ref. [3]. To do this one identifies l_μ with Σn_μ around the plaquette dual to the link on which l_μ lies; l_μ is therefore just the local vortex density and being a lattice curl, the condition $\nabla \cdot l = 0$ is obvious.

Rather than proceed directly with eq. (2.8) we will now show that a lattice model of scalar QED yields an expression closely related to it. To this end consider the model,

$$Z(\gamma, e) = \int_{-\pi}^{+\pi} \mathrm{d}\left[\frac{\varphi(i)}{2\pi}\right] \int_{-\infty}^{+\infty} \mathrm{d}[\chi_\mu] \times \sum_{\{n_\mu\}} \exp \sum \left[-\tfrac{1}{2}\gamma(\Delta_\mu \varphi - 2\pi n_\mu - e\chi_\mu)^2 - \tfrac{1}{2}(\Delta \wedge \chi)^2\right] . \tag{2.11}$$

This is a natural model for non-linear scalar QED with a *non-compact* gauge group (i.e., we do not identify rotations of 2π with the identity so that the group

$\mathbb{R}$ rather than U(1) or O(2)). Using the Poisson formula yields

$$Z \propto \int_{-\infty}^{+\infty} d[\chi_\mu] \sum_{n_\mu} \delta_{\Delta \cdot n, 0} \exp\left[\sum_{i,\mu} -\frac{n^2}{2\gamma} + ie\chi_\mu n_\mu - \tfrac{1}{2}(\Delta \wedge \chi)^2\right] . \tag{2.12}$$

This is the same as (2.8) as far as the topological excitations are concerned, provided we set $1/\gamma$ to zero and $2\pi\sqrt{\beta} = e$. If $\gamma^{-1} \neq 0$ we have added a chemical potential/unit length for the vortices. γ^{-1} therefore plays a role analogous to the $\ln y$ of ref. [3]. What has happened is that we have identified the vortices with the "particle" trajectories of the scalar QED. Since scalar QED also possesses topological excitations in the form of Nielson-Oleson vortices (provided that we are in a type II superconducting regime, non-linear QED is an extreme type II superconductor), we can continue from (2.12) in exact analogy with the progression (2.2)–(2.5), exhibit these vortices and identify them with the particle content of the original $O(2)_3$ model.

In (2.12) we put $n = \Delta \wedge M$, $M \in \mathbb{Z}$ and use the Poisson formula again to replace the sum over M by an integration over Φ_μ.

$$Z = N \int_{-\infty}^{+\infty} d[\chi_\mu] \int_{-\infty}^{+\infty} d[\Phi_\mu] \sum_{\{m_\mu\}} \exp\left[\sum -\frac{(\Delta \wedge \Phi)^2}{2\gamma}\right.$$
$$\left. - ie(\Delta \wedge \Phi) \cdot \chi - \tfrac{1}{2}(\Delta_1 \chi)^2 + 2\pi i m\Phi \right] . \tag{2.13}$$

Using Abels resummation formula (the discrete equivalent of integrating by parts), we transfer the lattice curl from Φ to χ in the second term of the exponent and rewrite

$$Z = N \int_{-\infty}^{+\infty} d[\chi] \int_{-\infty}^{+\infty} d[\Phi] \sum_{\{m_\mu\}} \exp[\sum -(\Delta \wedge \Phi)^2/2\gamma$$
$$- \tfrac{1}{2}(\Delta \wedge \chi + ie\Phi)^2 - \tfrac{1}{2}e^2\Phi^2 + 2\pi i m\Phi] . \tag{2.14}$$

We can now perform the χ integration to yield

$$Z = N \int d[\Phi] \sum_{m_\mu} \delta_{\Delta \cdot m, 0}$$
$$\times \exp[\sum_i - (\Delta \wedge \Phi)^2/2\gamma - \tfrac{1}{2}e^2\Phi^2 + 2\pi i m\Phi] . \tag{2.15}$$

The constraint on m again arises from gauge invariance. The m therefore form closed loops interacting with a Biot-Savart like force but with range $(e^2\gamma)^{-1}$. The original loops l_μ were sources of χ_μ (analogous to A_μ) and the new loops are sources of Φ_μ. (analogous to H_μ). The Gaussian integral relates them as

$$\Phi_\mu = \frac{i}{e}(\nabla \wedge \chi)_\mu \, .$$

These new m_μ loops are therefore Nielson-Oleson vortices. If we set $\gamma^{-1} = 0$ and perform the Φ integration we obtain

$$Z(\infty, e) = N \sum_m \delta_{\Delta \cdot m, 0} \exp\left(-\frac{4\pi^2}{e^2} \sum m_\mu^2\right) , \tag{2.16}$$

which is identical to (2.5) provided $2\pi\sqrt{\beta} = e$, the identification we have already made. Therefore, as we said earlier, the loops in (2.5) are the extreme short-range interaction limit of the Nielson-Oleson vortices. When $\gamma^{-1} \neq 0$ the model which is dual to the QED theory is (2.2) with the addition of a term in the exponent i.e.,

$$Z(\gamma, e) \propto \int \mathrm{d}\left[\frac{\theta_i}{2\pi}\right] \sum_{n_\mu} \exp\left[-\sum_i \{\tfrac{1}{2}\beta(\Delta_\mu \theta_i + 2\pi n_\mu(i))^2 + \tfrac{1}{2}\gamma(\Delta \wedge n)^2\}\right] , \tag{2.17}$$

or

$$Z(\gamma, e) \propto \int \mathrm{d}\left[\frac{\theta_i}{2\pi}\right] \sum_{n_\mu} \exp \sum_i \, [\tfrac{1}{2}\beta(\Delta_\mu \theta_i + 2\pi n_\mu(i))^2 - \left(\sum_{\text{plaquettes}} n\right)^2 / 2\gamma] \, . \tag{2.18}$$

What we have shown is that a theory with gauge group $\mathbb{Z}$ (i.e. 2.17) is dual to a theory with gauge group $\mathbb{R}$, with the topological excitations of one theory appearing as the particle trajectories in the other. We have been rather cavalier about throwing away the spin-wave contributions but in these theories they are clearly not relevant to any phase transitions.

3. Continuum approximation to the duality map

In this section we will set up a formalism for studying the statistical mechanics of closed-loop configurations. This formalism is simply the connection between the particle and field content of a quantum field theory. This duality is usually only exhibited by studying quantum field theory in the language of Hilbert spaces and representations of canonical commutation operators in terms of annihilation and creation operators. It is in the spirit of the times, however, to demonstrate that anything Hilbert space can do then path integrals can (with care!) do also.

Let $P(t)$ be the total number of closed, non-oriented, paths of total length t steps. We wish to compute:

$$\int \mathrm{d}t \, \mathrm{e}^{-At} P(t) \, . \tag{3.1}$$

We will use continuum notation because most of what we wish to study will take place on a length scale much greater than that of the lattice.

Introduce a quantity $p(x, x', t)$ as the probability that a random walk of t steps starting at x will arrive at x'. It is easy to see that $p(x, x', t)$ satisfies the equation

$$\Delta_t p = \frac{L^2}{2D} \Delta^2 p \tag{3.2}$$

provided $|x - x'| \ll Lt$. The solutions of this are well approximated by solutions of the continuum equation:

$$\frac{\partial p}{\partial t} = \frac{L^2}{2D} \nabla^2 p \,, \tag{3.2a}$$

and we will work with this equation from now on.

The solution to (3.2) can be written as a Gaussian path integral [11].

$$p(x, x', t) = N \int_{xt}^{x't'} \mathrm{d}[x] \exp\left\{-\int \left[\frac{D}{2L^2} \dot{x}^2\right] \mathrm{d}t\right\} \tag{3.3}$$

the normalization N being set so that $\int p = 1$.

The total number of paths from x, to x' in t steps is:

$$\begin{aligned} p(x, x', t)(2D)^t &= \Gamma(x, x', t) \\ &= N \int \mathrm{d}[x] \exp\left[-\int \left\{\frac{D}{2L^2} \; \dot{x}^2 - \ln 2D\right\} \mathrm{d}t\right] . \end{aligned} \tag{3.4}$$

Also

$$\int_0^\infty \Gamma(x, x, t)\, \mathrm{e}^{-At}\, \mathrm{d}t = \int \mathrm{d}t\, \mathrm{e}^{-At} P(t)\, 2t \,, \tag{3.5}$$

the factor 2 coming from the fact that a loop from x to x' is indistinguishable from a loop from x' to x and the factor t arises because x can be at t locations on the loop.

Writing

$$\begin{gathered} \Gamma(x, x', t) = \sum_n \Phi(x)\, \Phi(x')\, \mathrm{e}^{-\omega_n^2 |t|} \,, \\ \left\{-\frac{L^2}{2D} \nabla^2 - \ln 2D\right\} \Phi_n = \omega_n^2 \Phi_n \,, \qquad \int \Phi_n^2 = 1 \,, \end{gathered} \tag{3.6}$$

we see that

$$\int_0^\infty 2t\, P(t)\, \mathrm{e}^{-At} \mathrm{d}t = \sum_n \frac{1}{\omega_n^2 + A} = \mathrm{Tr}\left\{-\frac{L^2}{2D} \; \nabla^2 + A - \ln 2D\right\}^{-1} ,$$

so finally,

$$\int_0^\infty P(t)\, \mathrm{e}^{-At}\, \mathrm{d}t = -\tfrac{1}{2} \ln \det \left\{ -\frac{L^2 \nabla^2}{2D} + A - \ln 2D \right\} . \tag{3.7}$$

We have therefore managed to express the statistical sum over configurations of action A/unit step in terms of the well-known expression for a one-loop Feynman diagram. If we consider a non-interacting gas of such loops then we have to exponentiate the configuration sum for one loop:

$$\begin{aligned} Z &= \sum_n \{ \int_0^\infty P(t)\, \mathrm{e}^{-At}\, \mathrm{d}t \}^n / n! \\ &= \det{}^{-1/2} \left\{ -\frac{L^2 \nabla^2}{2D} + A - \ln 2D \right\} \\ &= \int \mathrm{d}[\varphi] \exp\left[-\int \mathrm{d}^D x \left\{ \tfrac{1}{2}\varphi \left\{ -\frac{L^2 \nabla^2}{2D} + A - \ln(2D) \right\} \varphi \right\} \right] . \end{aligned} \tag{3.8}$$

A gas of non-interacting loops with a certain energy or action per unit length can therefore be written as a free field theory with mass^2 = $A - \ln(2D)$.

The representation of the quantity Γ as a path integral enables one to introduce interactions between different elements of the loop. Consider firstly the interaction between a loop and an external magnetic field:

$$G(x, x', t) = \int_x^{x'} \mathrm{d}(x) \exp\, [-\int_0^t \{\tfrac{1}{2}\dot{x}^2 + V(x) + ie\dot{x}_\mu A_\mu\}\, \mathrm{d}t] . \tag{3.9}$$

G satisfies the equation

$$\frac{\partial G}{\partial t} = \tfrac{1}{2}(\partial_\mu + ieA_\mu)^2 G - VG , \tag{3.10}$$

enabling one to go through the previous analysis but with ∇ replaced by the covariant derivative. We can write the equation for a gas of loops interacting through Biot-Savart like forces (eq. (2.8)) between elements $\mathrm{d}x_1\, \mathrm{d}x_2$

$$V_{\text{int}} = \mathrm{d}x_1^\mu\, G_{\mu\nu}(x_1 x_2)\, \mathrm{d}x_2^\nu , \tag{3.11}$$

in the form

$$Z = N \int \mathrm{d}[\varphi]\, \mathrm{d}[\varphi^*]\, \mathrm{d}[A_\mu] \exp[-\int \mathrm{d}^D x\, \{\nabla\varphi\nabla\varphi^* + m^2 \varphi^*\varphi - \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}\}] . \tag{3.12}$$

The pair of complex fields is needed because now the orientation of loops is important and the factor $\frac{1}{2}$ in eq. (3.7) must vanish. We have exhibited once again, therefore, the connection between our generalized $O(2)_3$ model and scalar QED., this time in a way that seems valid in the limit $\gamma \to \infty$. If $m^2 < 0$, i.e., the free energy per unit length is negative, then the vacuum will be filled with a spaghetti of tangled loops and the equivalent field theory will be in the Higgs-Kibble or superconducting phase. In order to stop the density being infinite we need a $\lambda\varphi^4$ (with positive λ) to produce a short-distance repulsion in eq. (3.12). Such a term would always be generated in any renormalisation scheme for (3.12). In fact it should be necessary to introduce a bare $\lambda\varphi^4$ term in order to 'stiffen' the random walks on the lattice because eq. (3.2) allows paths which do not exist in the partition function we are considering. These paths are the backtracking or overlapping paths where the total charge along parts of them are zero. In the original partition function these have weight corresponding to $l_\mu = 0$ whereas in the path integral the weight is greater because of the possibility of backtracking. The $\lambda\varphi^4$ repels the paths and so reduces this effect.

It is worth explaining what the order parameter $\langle\varphi\rangle$ means in (3.12). Clearly if $m^2 > 0$ we expect $\langle\varphi\rangle = 0$ but $\langle\varphi^2\rangle \neq 0$, while for $m^2 < 0$ we expect to have broken U(1) symmetry with $\langle\varphi\rangle \neq 0$.

If we write a modified (3.8) with a scalar external interaction for the threads we find:

$$Z = \sum_n \frac{1}{n!} \{\int \mathrm{d}[x] \exp [-\int\{\tfrac{1}{2}\dot{x}^2 + \tfrac{1}{2}V(x)\}\, \mathrm{d}t]\}^n$$

$$= N\int \mathrm{d}[\varphi] \exp[-\int\{\varphi(-\tfrac{1}{2}\nabla^2)\varphi + \tfrac{1}{2}V(x)\varphi^2\}\, \mathrm{d}^D x]\ .$$

Differentiating this functionally with respect to V yields:

$$\frac{1}{Z}\frac{\delta Z}{\delta V(x)} = \frac{1}{2Z}\sum\int [\mathrm{d}x_1] \ldots [\mathrm{d}x_n] \frac{1}{n!} [\int \mathrm{d}t_1\, \delta(x - x_1) + \int \mathrm{d}t_2 (\delta(x - x_2) \ldots$$

$$+ \int \mathrm{d}t_n\, \delta(x - x_n)] [\exp - \int\{\tfrac{1}{2}\dot{x}^2 + \tfrac{1}{2}V(x)\}\mathrm{d}t]^n$$

$$= \tfrac{1}{2}\langle\rho(x)\rangle = \langle\tfrac{1}{2}\varphi^2(x)\rangle\ ,$$

where ρ is the local density of threads.

Similar thought about $\langle\varphi\rangle$ indicates that it is that proportion of the threads which is infinite in length. Fortunately the correlation inequality

$$\langle\varphi^2\rangle \geqslant \langle\varphi\rangle\langle\varphi\rangle$$

shows that this fraction is less than the total number of threads!

4. Renormalization group and discussion

Having decided which field theory describes the problem in hand the logical thing to do is investigate the behaviour as renormalised m^2 goes through zero by applying the

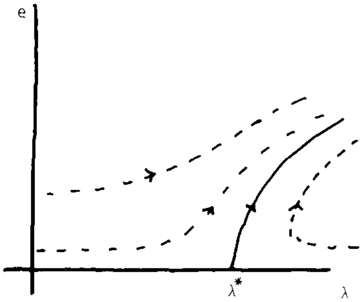

Fig. 2. The flow with changing scale in the coupling constant plane, showing the instability in the e direction at the critical point.

renormalization group in $4 - \epsilon$ dimensions. The one-loop β (e, λ) function is easily calculated in the ϵ expansion and is

$$\beta_\lambda = -\epsilon\lambda + [10\lambda^2 + 36e^4 - 24e^2\lambda]/16\pi^2 \ ,$$

$$\beta_e = -\epsilon e + [2e^3/3]/16\pi^2 \ , \tag{4.1}$$

This has a fixed point at $\lambda^* = \frac{8}{5}\pi^2$, $e^* = 0$ which is the usual $4 - \epsilon$ fixed point. If this dominated the phase transition then we would get the usual exponents (to all orders) that one would get from the O(2) model. Together with the usual symmetry of exponents (where defined) above and below the critical point we would have found consistency in our duality transformation. Unfortunately the fixed point is unstable in the e direction and there are no other fixed points nearby (fig. 2). Halperin, Lubensky and Ma [4] have argued that the phase transition is in fact first order.

We have to understand why this is so and not necessarily inconsistent with our usual understanding of phase transitions in O(N) symmetric systems.

First-order phase transitions are usually co-operative phenomena and the nature of the co-operation is not hard to find. Naively the phase transition occurs when the free energy/unit length of our thread goes negative. The energy part of the free energy is just the energy locked up in the magnetic field of a long wire. In three dimensions this is logarithmically infrared divergent and certainly depends on the density of other currents. In general, neighbouring currents will be oppositely oriented thereby reducing the free energy of the system. The dependence of the thread concentration will be non-linear causing the effective potential to have the form of fig. 3. This mechanism is identical to that which happens to the free energy of dislocations in a solid near its melting point (see ref. [6]) although the interactions between dislocations would need mixtures of higher-spin fields to describe their behaviour correctly. It is well-known that melting of solids is a first-order phase transition.

Is it possible for the non-linear O(2)$_3$ model to have a first-order transition? For the generalized model of sect. 2 the forces between the Nielsen-Olesen vortices

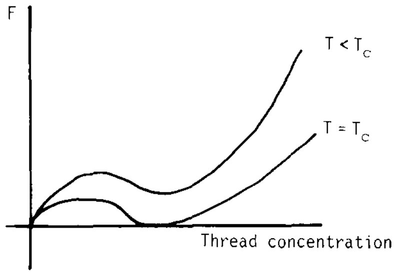

Fig. 3. The total free energy of the threads as a function of concentration.

behave in a similar manner below the transition point so the theory seems self consistent. One would however have expected the $O(2)_3$ model to be in the same universality class as the linear $O(2)_3$ model and therefore have a second order phase transition described by well known critical exponents in $(4 - \epsilon)$. Certainly the $O(N)_3$ non-linear models seem to be for $N \geqslant 3$. One can carry out expansions in $2 + \epsilon$ for these non-linear models [12] and find the same qualitative agreement. For $N = 2$, the $2 + \epsilon$ expansions do not work because the naive continuum model is simply a free theory and always seems to be in the ordered phase with non-vanishing value for the magnetisation. The phase transition must be dominated by the topological excitations therefore as is the $O(2)_2$ model. This is possibly the escape from paradox. The linear $O(2)_3$ model behaves, in the $4 - \epsilon$ expansion, no differently from the other $O(N)$ models and its phase transition is therefore dominated by "spin wave" type fluctuations as are all the other cases. This option is not available to the non-linear case. There are presumably topological excitations in the linear model in the continuum limit but they do not seem to play any significant role.

After submitting this paper we received a preprint by Peskin [13] discussing the same topic. Our duality map (sect. 2) is identical with his, while our sect. 3 contains matter not discussed by him. Peskin assumes that the phase transition is second order in the case $\gamma = 0$. If this is true (and there is evidence from high-temperature series [14]) then γ has to be a relevant parameter in the sense of the renormalization group.

We wish to thank David Wallace, Michael Gunn and Mike Kosterlitz for useful conversations. We gratefully acknowledge the financial support of the Science Research Council.

References

[1] T. Banks, R. Myerson and J. Kogut, Nucl. Phys. B124 (1977) 499.
[2] J.M. Kosterliz and D.J. Thouless, J. Phys. C6 (1973) 1181.
[3] J.V. Jose, L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, Phys. Rev. B16 (1977) 1217.

[4] B.I. Halperin, T.C. Lubensky and Shang-keng Ma, Phys. Rev. Lett. 32 (1974) 292.
[5] C. Montonen and D. Olive, Phys. Lett. 72B (1977) 117;
S. Coleman, Phys. Rev. D11 (1975) 2088;
S. Mandelstam, Phys. Reports 23 (1976) 307.
[6] F.R.N. Nabarro, Theory of crystal dislocations (Oxford University Press, 1967).
[7] P.R. Thomas and M. Stone, DAMTP preprint.
[8] J. Villain, J. de Phys. 36 (1975) 581.
[9] M.B. Einhorn and R. Savit, Fermilab. preprints;
J. Glimm and A. Jaffe, Comm. Math. Phys. 56 (1977) 195.
[10] H.B. Nielsen and P. Olesen, Nucl. Phys. B61 (1973) 45.
[11] R. Feynman and A.R. Hibbs, Quantum mechanics and path integrals, (McGraw-Hill, 1965)).
[12] E. Brézin and J. Zinn-Justin, Phys. Rev. B14 (1976) 3110.
[13] M. Peskin, Harvard preprint HVTP 77/A083.
[14] J. Kogut, private communication.