PHYSICAL REVIEW D VOLUME 19, NUMBER 4 15 FEBRUARY 1979

Nonlinear two-dimensional σ-model instanton as a tunneling process

Khalil M. Bitar,* Shau-Jin Chang, Garland Grammer, Jr., and John D. Stack
Physics Department, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801
(Received 30 May 1978)

We study the instanton in the nonlinear two-dimensional σ model and show that it can be interpreted, in Minkowski space, as a tunneling process through a potential barrier between two vacuums. In this case the process carries nontrivial winding number. We then show, using this interpretation, that the σ-model vacuum is nevertheless unique by demonstrating that two such vacuums may also be connected by processes that carry zero winding number and which do not require tunneling through a barrier. Some geometrical aspects of instanton solutions are also given.

I. INTRODUCTION

It is well known that the nonlinear σ model in two dimensions shares many properties with non-Abelian gauge theories in four dimensions. In particular, both theories are asumptotically free, and both contain instantion or pseudoparticle solutions of the field equations. Nevertheless the nonlinear σ model is a much simpler theory than a non-Abelian gauge theory, first because it is two dimensional, but second and more important, its global rotational invariance is enormously simpler than the local invariance of a gauge theory.

In this paper, we wish to discuss certain aspects of instanton physics in the $n=3$ nonlinear σ model. As the title implies, tunneling is the main focus. The σ model provides a simple example in which to study the process of tunneling through an instanton directly in Minkowski space. A formalism for doing this has recently been developed by two of us and applied to the instanton in gauge theories.[1,2] The calculation of the tunneling amplitude as well as the way the angles defining the σ field move around on the unit sphere during the tunneling process are discussed in some detail in the body of the paper in Secs. III and IV.

In addition, we have been interested in comparing and contrasting the vacuum structure and the role played by instantons in determining the properties of the vacuum in the two cases. The existence of a topological charge, ususally called winding number. The winding number q is the integral of a density ρ over (Euclidean) space-time. The instanton solutions carry definite, integer values of q, so winding number is first associated with the whole tunneling process which begins and ends in a classical vacuum configuration. However, in gauge theories in certain gauges, e.g. the temporal gauge, winding number also becomes a label for classical vacuum configurations, which involves properties of the system at a fixed time. This arises because the winding number density ρ can be expressed as a total divergence of a current K_μ and surface terms at spatial infinity are dropped. Thus q, by Gauss's law, is expressible in terms of the difference at large positive and negative times of the spatial integral of the time component of K_μ. This allows the winding number label to be attached to the classical vacuum configurations themselves and leads to the multiple- or θ-vacuum description of the gauge theory.

Since the nonlinear σ model shares so many properties with gauge theories, it is fair to ask if a multiple-vacuum description is possible here too. As is clear from the above discussion, a necessary prerequisite is that the winding number density ρ be expressible as the total divergence of a current K_μ. The σ model does have this feature as discussed in Sec. III A. However, upon applying Gauss's law, the contributions from the surface at spatial infinity are not negligible in general, and in fact, can be as important as the contributions of the infinite time surfaces. This evidently prevents the assignment of any physically meaningful topological index to the classical vacuum configurations themselves. In Sec. V we argue that there is no multiple-vacuum description for the nonlinear σ model.

It should perhaps be pointed out that the notion of tunneling is still meaningful even in the absence of a multiple-vacuum description. The quantum theory of a simple pendulum provides an example. A unique classical vacuum configuration can be defined, but nevertheless in the quantum theory tunnelings in which the pendulum turns all the way over and returns to this classical vacuum configuration will occur. These tunnelings can play an important role in understanding the quantum ground state if the gravitational field is sufficiently weak. Our point here is that one need not tunnel from a given field configuration to a different one, but that initial and final configurations can be the same if the field space is multiply connected.

The paper is arranged as follows. In Sec. II we review the Minkowski description of tunneling. Section III A covers the basic properties of the σ model. The pseudoparticle solutions found by Polyakov[3] are discussed and the current, K_μ, whose divergence is the winding number density, is identified. In Sec. III B we discuss the geometrical interpretation of the model while Sec. IV contains the actual calculation of the tunneling amplitude. Finally in Sec. V we address the question of multiple vacuums in the σ model.

II. TUNNELING IN MINKOWSKI SPACE

The one-dimensional quantum-mechanical tunneling problem involves finding the wave function at some point b, given the wave function at another point a, when a and b are separated by a classically forbidden region. There are, of course, many ways of calculating this wave function. In the WKB approach, the amplitude for a particle of energy E to tunnel through a slowly varying potential V is e^{-R_0} where the function R_0 satisfied in the (zeroth-order) WKB approximation

$$(\nabla_x R_0)^2 = 2m[V(x) - E] \tag{2.1}$$

or

$$R_0 = \int_a^b dx\{2m[V(x) - E]\}^{1/2}, \tag{2.2}$$

where m is the mass of the particle. The wave function at point b is usually written as

$$\psi_b = e^{iS_0}\psi_a, \tag{2.3}$$

where $S_0 = iR_0$. Higher-order corrections can also be computed systematically, assuming that all of the conditions for the validity of the WKB method are satisfied. A similar result emerges from the Feynman path-integral approach.

The multidimensional tunneling problem can also be solved in the WKB approximation.[1,2] The tunneling amplitude follows from the obvious generalization of Eq. (2.1). This differential equation is intractable, however, unless we can reduce it to an equivalent one-dimensional problem. This is done as follows. Consider a path in the multidimensional space connecting the points of interest, a and b. At each point l on the path, choose a local orthogonal coordinate system in the field space with one axis tangent to the path at l; the rest, call them collectively $\vec{n}$, will be perpendicular. If a path can be found for which the derivatives of R in all of the n_i directions vanish, then only variations along l need be considered and the problem becomes one dimensional. This particular path is called a most probable escape path (MPEP) and lies along a minimum of R. Should there be several such paths, the corresponding contributions to the wave function would be added. Contributions to the amplitude from other paths will be exponentially small in comparison.

For convenience we will parametrize the MPEP as $l(\lambda)$ with $\lambda_0 \leq \lambda \leq \lambda_1$ such that $\lambda = \lambda_0$ corresponds to the point a and $\lambda = \lambda_1$ to b. The tunneling amplitude is then simply e^{-R_0} with

$$R_0 = \int_{\lambda_0}^{\lambda_1} d\lambda\{2m(\lambda)[V(\lambda) - E]\}^{1/2} \tag{2.4}$$

and $S_0 = iR_0$. Note that in general the effective mass of the particle becomes λ dependent.

Consider, for a moment, an arbitrary path that does not correspond to a minimum of R, that is, not an MPEP. At a given point on that path, linear variations in the n_i direction need not vanish and thus the problem does not reduce to a one-dimensional equation. In the local coordinate system, Eq. (2.1) becomes

$$\begin{aligned}(\nabla_l R)^2 &= 2m(\lambda)(V - E) - (\nabla_\perp R)^2 \\ &\leq 2m(\lambda)(V - E).\end{aligned} \tag{2.5}$$

If variations in all but the l direction are ignored, the solution obtained will constitute an upper bound on R according to Eq. (2.5), and thus will give a lower bound on the contribution of that path to the amplitude e^{-R}. If one then minimizes this R, the MPEP contribution is again obtained. A complete discussion of this method is given in Refs. 1 and 2.

After discussing the nonlinear σ model and its instanton solutions in the next section, we will use this Minkowski-space tunneling formalism to discuss the instanton as a tunneling path in the (infinite dimensional) field space.

III. THE NONLINEAR σ MODEL

A. The Euclidean solutions

We work in two dimensions with a single three-component scalar field $\vec{s}(\vec{x}) = \vec{\sigma}(x)/g$ of constant length

$$\vec{s}^{\,2}(\vec{x}) = 1/g^2, \tag{3.1}$$

or $\vec{\sigma}^2 = 1$. The 1 axis is the ordinary space direction, while the Euclidean time direction is the 2 axis. Because of Eq. (3.1), $\vec{s}$ can be parametrized in terms of two scalar fields θ and Φ according to

$$\vec{s}(\vec{x}) = \frac{1}{g}(\cos\theta(\vec{x}), \sin\theta(\vec{x})\cos\Phi(\vec{x}), \sin\theta(\vec{x})\sin\Phi(\vec{x})) \tag{3.2}$$

and additionally we require that

$$\vec{s}(\vec{x}) = \frac{1}{g}(1, 0, 0), \quad r \to \infty \tag{3.3}$$

that is, $\theta \to 0$ at spatial infinity, $r=(x_1^2+x_2^2)^{1/2} \to \infty$. We will use the term "classical vacuum configuration" to refer to a situation in which at a fixed x_2 $\vec{s}(x_1,x_2)=(1/g)(1,0,0)$ for all x_1.

The two-dimensional field theory is then defined by the Lagrangian density

$$\mathcal{L}(\vec{x})=\tfrac{1}{2}[\nabla_\mu s^a(\vec{x})]^2$$

$$=\frac{1}{2g^2}\{[\nabla_\mu\theta(\vec{x})]^2+\sin^2\theta(\vec{x})[\nabla_\mu\Phi(\vec{x})]^2\}, \quad (3.4)$$

where $\mu=1,2$. For later convenience we define the two-component vectors

$$a(\vec{x})=(\nabla_1\theta(\vec{x}),\sin\theta(\vec{x})\nabla_1\Phi(\vec{x})) \quad (3.5)$$

and

$$b(\vec{x})=(\nabla_2\theta(\vec{x}),\sin\theta(\vec{x})\nabla_2\Phi(\vec{x})) \quad (3.6)$$

and in terms of which

$$\mathcal{L}(\vec{x})=\frac{1}{2g^2}[a^2(\vec{x})+b^2(\vec{x})]. \quad (3.7)$$

The x_1-x_2 plane $\tilde{S}^2$ is equivalent to a two-sphere S^2 if all points at $r=\infty$ are identified as in Eq. (3.3). The field $\vec{s}(x)$ then constitutes a mapping of $S^2\to\tilde{S}^2$, with the degree of mapping given by

$$q=\frac{g^3}{8\pi}\int d^2x\,\epsilon_{abc}\epsilon_{\mu\nu}s^a\nabla_\nu s^b\nabla_\mu s^c$$

$$=\frac{1}{4\pi}\int d^2x\,\epsilon_{\mu\nu}a^\mu(\vec{x})b^\nu(\vec{x})$$

$$=\frac{1}{4\pi}a\times b\,. \quad (3.8)$$

For the dot product we write

$$a\circ b=\int d^2x\,a(\vec{x})\cdot b(\vec{x}) \quad (3.9)$$

and for the cross product

$$a\times b=\int d^2x\,\epsilon_{\mu\nu}a^\mu(\vec{x})b^\nu(\vec{x})\,. \quad (3.10)$$

The winding number q assumes only integer values for fields satisfying the condition (3.3). It can be used to label the different classes of finite-action field configurations as we will discuss in a moment.

The action is bounded by

$$S=\frac{1}{2g^2}(a^2+b^2) \quad (3.11)$$

$$\geq |a|\,|b|/g^2 \quad (3.12)$$

$$\geq|a\times b|/g^2=\frac{4\pi q}{g^2}, \quad (3.13)$$

where $|a|=(a^2)^{1/2}$ using Eq. (3.9). The minimum action for each topological class thus occurs for field configurations which satisfy the two conditions

$$|a|=|b| \quad (3.14)$$

and

$$a\circ b=0\,. \quad (3.15)$$

We will find that the explicit solutions, given below, satisfy Eqs. (3.14) and (3.15) even in the unintegrated form with $a(\vec{x})$ and $b(\vec{x})$.

The most general solutions of the equations of motion which follow from the Lagrangian, Eq. (3.4), were first given by Polyakov,[3] and are

$$\cot\frac{\theta(\vec{x})}{2}e^{i\Phi(\vec{x})}=\prod_i\left(\frac{z-z_i}{c}\right)^{m_i}\prod_j\left(\frac{c}{z-z_j}\right)^{n_j},$$

where $z=x_1+i\,x_2$ and c, the "size" of the instanton, is a constant. The winding number is $q=\sum m_i>\sum n_j$.

Consider now the simplest nontrivial case of winding number one, that is, $m_1=1$, $m_i=0$, $i>1$, $n_j=0$. The spin field is given by

$$\theta(\vec{x})=\cos^{-1}\left(\frac{r^2-c^2}{r^2+c^2}\right) \quad (3.16)$$

and

$$\Phi(\vec{x})=\tan^{-1}\left(\frac{x_2}{x_1}\right), \quad (3.17)$$

where $r=(x_1^2+x_2^2)^{1/2}$ and we have chosen the position of the instanton $z_1=0$ for convenience. The constant c determines the "size" of the instanton. Equations (3.5) and (3.6) then give

$$a(\vec{x})=\frac{2c}{r(r^2+c^2)}(-x_1,x_2) \quad (3.18)$$

and

$$b(\vec{x})=\frac{-2c}{r(r^2+c^2)}(x_2,x_1), \quad (3.19)$$

which satisfy $|a(\vec{x})|=|b(\vec{x})|$ and $a(\vec{x})\circ b(\vec{x})=0$ as we promised earlier. From Eq. (3.11) we find that $S=4\pi/g^2$ which is what we expected for the minimum-action solution belonging to the class $q=1$.

Finally we wish to identify the current K_μ whose divergence gives the winding number density ρ. Defining ρ to be the integrand of Eq. (3.8), i.e.,

$$q=\int\rho(x)d^2x, \quad (3.20)$$

we find from (3.8)

$$\rho=\frac{1}{4\pi}(\partial_1\cos\theta\partial_2\Phi-\partial_2\cos\theta\partial_1\Phi)\,. \quad (3.21)$$

We can express ρ as the divergence of a current K_μ given by

$$K_\mu = \frac{1}{4\pi}\epsilon_{\mu\nu}(1+\cos\theta)\partial_\nu\Phi , \qquad (3.22)$$

where K_μ satisfies

$$\partial_\mu K_\mu = \rho . \qquad (3.23)$$

In (3.22), we have included a divergenceless term $(1/4\pi)\epsilon_{\mu\nu}\partial_\nu\Phi$ in the definition of K_μ. The modification is necessary to remove the singularity associated with the polar coordinates at $\theta=\pi$.

B. Geometrical interpretation of the model

Since the isovector field $\vec{s}$ has fixed length $1/g$, its components may be associated with the coordinates of a point on the surface of a sphere of radius g^{-1}. According to Eq. (3.2) we take these to be the polar angle θ and the azimuthal angle Φ which are, of course, functions of x_1 and x_2. Thus $\theta(\vec{x})$ and $\Phi(\vec{x})$ constitute a mapping of the compactified x_1-x_2 plane onto the surface of the spheres where all points at infinity map onto the north pole.

The two-component vectors $\vec{a}$ and $\vec{b}$ defined in Eqs. (3.5) and (3.6) also have a simple interpretation. For motion of the point (x_1, x_2) in the x_1-x_2 plane at constant x_1, $\vec{a}$ is the corresponding "velocity vector" of the point (θ, Φ) on the surface of the sphere. Similarly, $\vec{b}$ is the "velocity vector" of the same point associated with the motion of the point (x_1, x_2) at constant x_2. In other words, $\vec{a}$ is the local tangent to the curve on the sphere that is the map of the straight line $x_1=$ constant and $\vec{b}$ is the local tangent for $x_2=$ constant.

Any infinite line in the x_1-x_2 plane has as its map a closed continuous curve on the surface of the sphere, passing through the pole ($\theta=0$ is always associated with $x_1\to\pm\infty$ or $x_2\to\pm\infty$). This is illustrated in Fig. 1. The fractional solid angle that a curve $x_2=$ const subtends at the origin of the sphere is given by

$$q(x_2)=\frac{1}{4\pi}\int_c d\phi(\vec{x})d\theta(\vec{x})\sin\theta(\vec{x}) , \qquad (3.24)$$

which is equivalent to Eq. (3.8) integrated only over the interval $-\infty \le x_2' \le x_2$.

The expression Eq. (3.24) is nothing but the winding number carried by the fields θ, Φ at x_2. At $x_2=-\infty$, the entire curve collapses onto the north pole corresponding to an aligned spin configuration.

At a later "time," $q(x_2)$ is nonzero and, in general, nonintegral according to Eq. (3.24). If as $x_2\to+\infty$ the curve sweeps the whole surface of the sphere once and then contracts to the north pole, then $q(\infty)=1$. If the sphere is covered n times as x_2 progresses from $-\infty$ to ∞, then $q(\infty)=n$ and one says that the mapping carries winding number n. In contrast, if the curve only sweeps part of the surface and then contracts back to the pole at $x_2\to+\infty$, the winding number does not change. As a mapping proceeds in x_2, it can be labeled at any time by $q(x_2)$ and the difference in winding number for field configurations at two different x_2 points can be determined by following q from $(x_2)_a$ to $(x_2)_b$.

The instanton that carries one unit of winding number is easily described in this geometrical picture. We found in the preceding section that the condition for minimum action is $\vec{a}\cdot\vec{b}=0$, that is, the "velocity vectors" on the sphere for motion along constant x_1 and along constant x_2 are orthogonal. Thus the grid of orthogonal constant x_1 and constant x_2 lines in the plane should map to a set of orthogonal curves on the surface of the sphere. An example of such a mapping is an inversion with center P (the pole of the sphere) and constant radius g^{-1}, since an inversion preserves angles. In this case the curves on the sphere are circles passing through P; the circles corresponding to constant x_1 are orthogonal to the circles for constant x_2 at their points of intersection. The angles θ and Φ can be easily computed in terms of $\vec{x}$ using Fig. 2 and are precisely the solutions found in Eqs. (3.16) and (3.17).

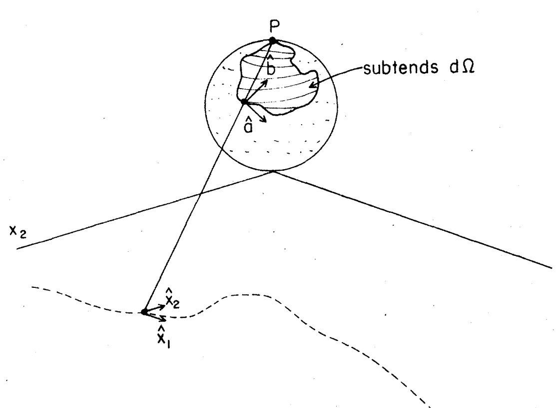


FIG. 1. The projection of the spin sphere onto the x_1-x_2 plane.

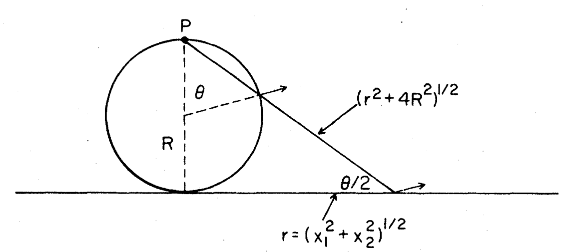


FIG. 2. The projection defining the angles of the $q=1$ instanton.

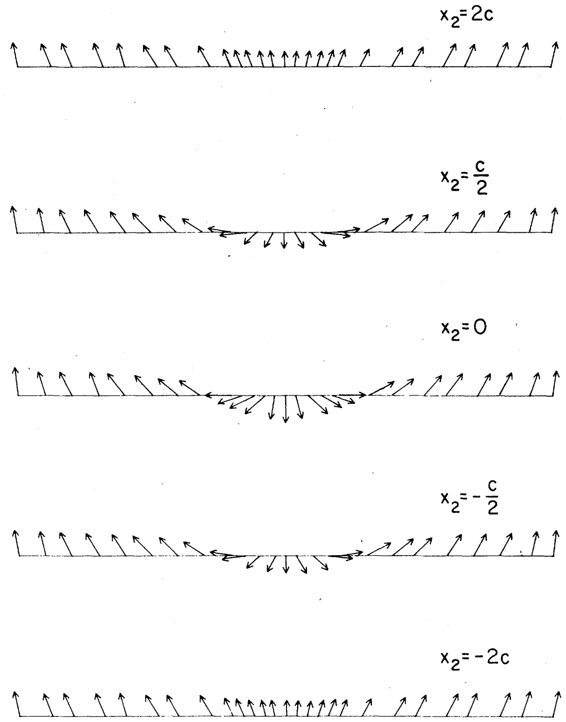


FIG. 3. The spin configuration for various values of x_2.

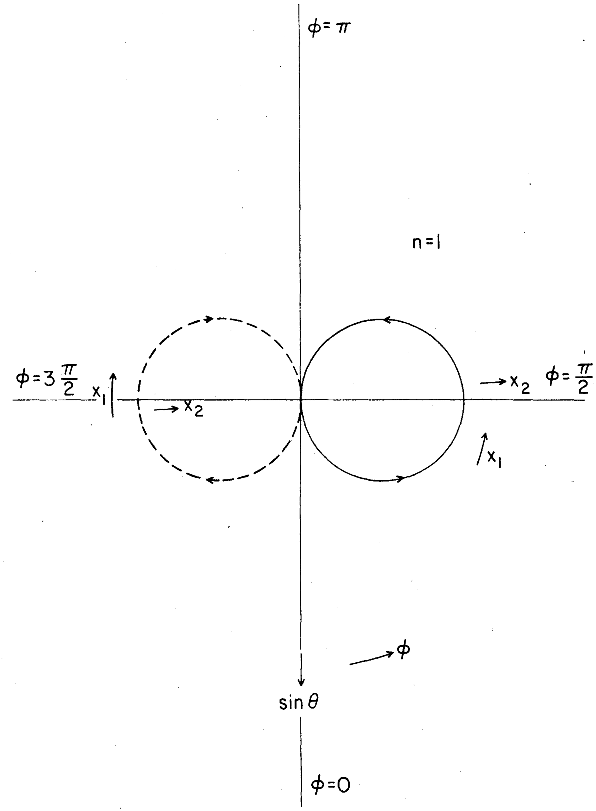


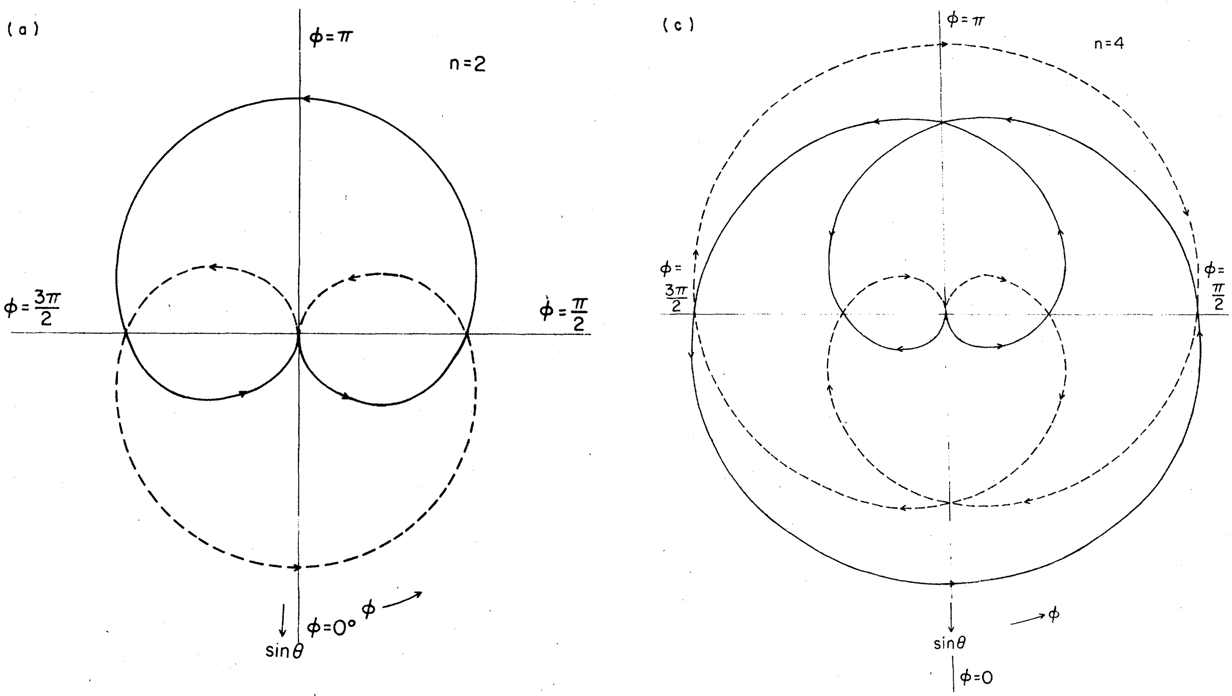
FIG. 4. The curves traced on the sphere for $x_2 = \pm T/2$, for the $q = 1$ instanton.

It is instructive to compare the development of the spin configurations in the $\vec{x}$ plane with the one-instanton solution θ and Φ on the sphere as a function of increasing time x_2 as shown in Fig. 3. We can think of the spin configuration at $x_2 = -\infty$ as being comprised of a row of atoms joined together by a string along the x_1 direction, with all of their spins aligned in the up direction. At a slightly later time, $x_2 = -T/2$, with T large, some disalignment has set in (smoothly) according to Eqs. (3.16) and (3.17), which then develops according to the later x_2 slices shown in the figure. Note that the spins near the ends of the string $x_1 = \pm\infty$ do not wind around the string as $x_2 \to +\infty$, but only dip slightly in θ as Φ progresses from 0 to 2π. Those spins near the constant x_1 line that passes through the south pole of the sphere do, however, dip below the string in θ and thus do wind completely around the string. At $x_2 = +\infty$ the spins are again all aligned in the up direction. The string connecting the atoms now has two "twists" in it because only a finite length of spins wrapped around while the others did not. In general, an instanton carrying integer winding number q connects spin-aligned states at $x_2 = \pm\infty$ which differ by $q+1$ "twists" in the string. The twists we discuss here are only a useful guide in visualizing the mapping of the sphere in a given winding number sector. As we discuss further in Sec. V, there is no physical difference between the aligned configurations at $x_2 = \pm\infty$. It is clear from observing the development in x_2 that the $q = 1$ instanton solution defines a mapping of the plane into the sphere once. For multi-instanton solutions tracing the x_2 development makes clear the q-fold nature of those mappings. The curves traced on the sphere for a large early time $x_2 = -T/2$ and a large late time $x_2 = +T/2$ are shown in Figs. 4 and Figs. 5(a)–5(c) for $q = 1, 2, 3, 4$.

IV. PSEUDOPARTICLE AS A TUNNELING PROCESS

A. Field configuration associated with an MPEP

We can now discuss the instanton solutions in Minkowski space as a tunneling between two aligned-spin configurations which takes place through a mapping with nonvanishing winding number. According to the formalism of Sec. II, the tunneling problem is first reduced to a one-dimensional quantum-mechanical problem in terms of a dynamical variable $\lambda \in [-\infty, \infty]$, where λ parametrizes the intermediate field configurations with $\lambda = -\infty$ corresponding to the initial aligned-



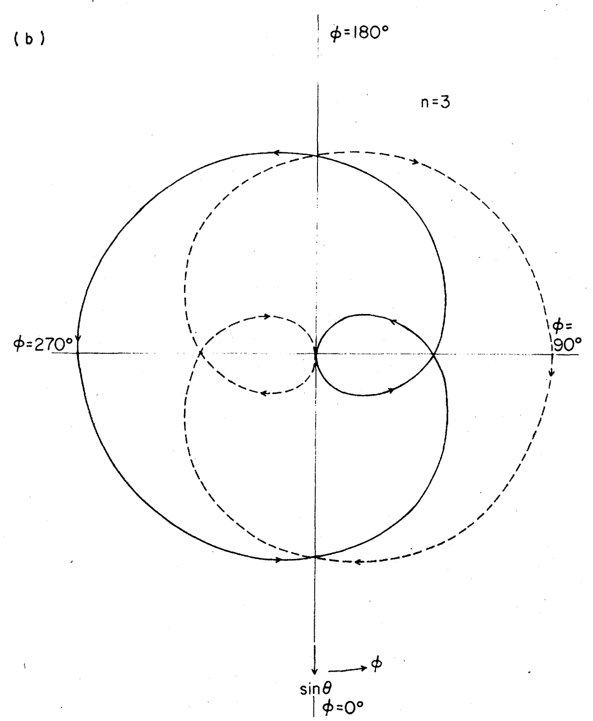


FIG. 5. The curves traced on the sphere for $x_2 = \pm T/2$, for (a) $q=2$, (b) $q=3$, (c) $q=4$.

spin state and $\lambda=\infty$ to the final state. The magnitude of the wave functional at each λ gives the relative probability for that configuration to occur; in particular, the transition amplitude is proportional to the ratio of the magnitude of the wave functional at $\lambda=-\infty$ and $\lambda=\infty$. The second step is to maximize the tunneling amplitude by determining the optimal intermediate field configuration, the most probable escape path (MPEP).

The Minkowski-space Lagrange function for the nonlinear σ model is

$$\mathcal{L}=\tfrac{1}{2}\{\mathring{\theta}^2-(\nabla\theta)^2+\sin^2\theta[\dot{\Phi}^2-(\nabla\Phi)^2]\}. \quad (4.1)$$

We will parametrize the intermediate field configuration according to

$$\theta=\theta(x,\lambda(t)) \quad (4.2)$$

and

$$\Phi=\Phi(x,\lambda(t)), \quad (4.3)$$

where θ and Φ correspond to aligned-spin configurations at $\lambda(t=-\infty)=-\infty$ and $\lambda(t=\infty)=\infty$; there is a change of winding number $q=1$ when λ varies from $-\infty$ to $+\infty$.

The analogs of Eqs. (3.5) and (3.6) are then ($x_1 \to x$ and $x_2 \to t$)

$$A(\vec{x})=(\nabla_x\theta(x,\lambda),\ \sin\theta(x,\lambda)\nabla_x\Phi(x,\lambda)) \quad (4.4)$$

and

$$B(\vec{x})=(\nabla_\lambda\theta(x,\lambda),\ \sin\theta(x,\lambda)\nabla_\lambda\Phi(x,\lambda)) \quad (4.5)$$

with the inner and cross products defined similarly to those of the preceding section. The winding number is generalized to describe the noninteger values of the intermediate field configuration according to

$$q(\lambda)=\frac{1}{4\pi}\int_{-\infty}^{\lambda}d\lambda'\int dx\,\epsilon_{\mu\nu}A^{\mu}(x,\lambda')B^{\nu}(x,\lambda').\quad(4.6)$$

The effective one-dimensional Lagrangian is

$$\mathcal{L}=\tfrac{1}{2}m(\lambda)\dot{\lambda}^2-V(\lambda),\quad(4.7)$$

where

$$m(\lambda)=\frac{1}{g^2}\int dx\,B^2(x,\lambda)\quad(4.8)$$

and

$$V(\lambda)=\frac{1}{2g^2}\int dx\,A^2(x,\lambda).\quad(4.9)$$

Using the WKB method, we can solve the single-parameter Langrangian Eq. (4.7) for the tunneling amplitude, and find an amplitude of e^{-R} with

$$R=\int_{-\infty}^{\infty}d\lambda[2m(\lambda)V(\lambda)]^{1/2}\quad(4.10)$$

$$=\frac{1}{g^2}\int_{-\infty}^{\infty}d\lambda\left[\int dx\,B^2(x,\lambda)\int dx\,A^2(x,\lambda)\right]^{1/2},\quad(4.11)$$

which is bounded by

$$R\geq\frac{1}{g^2}A\times B\quad(4.12)$$

$$=\frac{4\pi q}{g^2}.\quad(4.13)$$

The maximum rate corresponds to an intermediate field configuration which satisfies

$$A\cdot B=0,\quad(4.14)$$

although the condition $|A|=|B|$ is not required as it was in the Euclidean treatment.

An explicit parametrization can be obtained by replacing the time coordinate x_2 in the $q=1$ Euclidean pseudoparticle solution by λ:

$$\theta(x,\lambda(t))=\cos^{-1}\left(\frac{x^2+\lambda^2-c^2}{x^2+\lambda^2+c^2}\right)\quad(4.15)$$

and

$$\Phi(x,\lambda(t))=\tan^{-1}\left(\frac{\lambda}{x}\right).\quad(4.16)$$

According to Eqs. (3.16) and (3.17) these configurations provide a path between aligned states and carry winding number one. Although Eqs. (4.15) and (4.16) are simply related to the Euclidean solutions, we emphasize that λ is not a time parameter. In particular, it is not the Euclidean time nor an analytic continuation of it. In the WKB approach, λ labels the position along a most probable escape path, which as explained earlier, determines the main variation of the wave functional.

For the above configurations we have

$$A(x,\lambda)=\frac{2c}{r(r^2+c^2)}(-x,\lambda(t))\quad(4.17)$$

and

$$B(x,\lambda)=\frac{-2c}{r(r^2+c^2)}(\lambda(t),x)\quad(4.18)$$

with $r^2=x^2+\lambda^2$, for which $A\cdot B$ vanishes trivially since we had $a\cdot b=0$ for the corresponding Euclidean solutions. Actually, any multiple of the above A and B could be used as there is no restriction $|A|=|B|$ in the WKB development.

B. Tunneling potential in winding number space

The tunneling potential, according to Eq. (4.9), is

$$V(\lambda)=\tfrac{1}{2}m(\lambda)=\frac{\pi c^2}{g^2}\frac{1}{(\lambda^2+c^2)^{3/2}}.\quad(4.19)$$

The WKB tunneling amplitude e^{-R} is then from Eq. (4.11), simply

$$R=4\pi/g^2,\quad(4.20)$$

which is the appropriate rate for a unit winding number.

Since both V and q are given as a function of λ, the tunneling potential can be expressed as a function of q alone. Of course, the $q(\lambda)$ defined in Eq. (4.6) is far from unique; we can use any function $Q(\lambda)$ which agrees with q at integer values. In particular, we wish to rewrite the kinetic term in the Lagrangian so that

$$\tfrac{1}{2}m(\lambda)\dot{\lambda}^2\to\tfrac{1}{2}M\dot{Q}^2\quad(4.21)$$

with M independent of Q. The appropriate replacement is

$$Q(\lambda)=\frac{q\int_{-\infty}^{\lambda}d\lambda'[m(\lambda')]^{1/2}}{\int_{-\infty}^{\infty}d\lambda'[m(\lambda')]^{1/2}}\quad(4.22)$$

with

$$M=\left\{\frac{1}{q}\int_{-\infty}^{\infty}d\lambda'[m(\lambda')]^{1/2}\right\}^2.\quad(4.23)$$

For the explicit configurations of Eqs. (4.15) and (4.16), we have

$$Q(\lambda)=\frac{1}{2}\left[1+\mathrm{sgn}(\lambda)\frac{B_{\xi}(\frac{1}{2},\frac{1}{4})}{B(\frac{1}{4},\frac{1}{2})}\right]\quad(4.24)$$

and

$$M=\frac{2\pi c}{g^2}\left[\frac{\Gamma(\frac{1}{4})\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{4})}\right]^2,\quad(4.25)$$

where

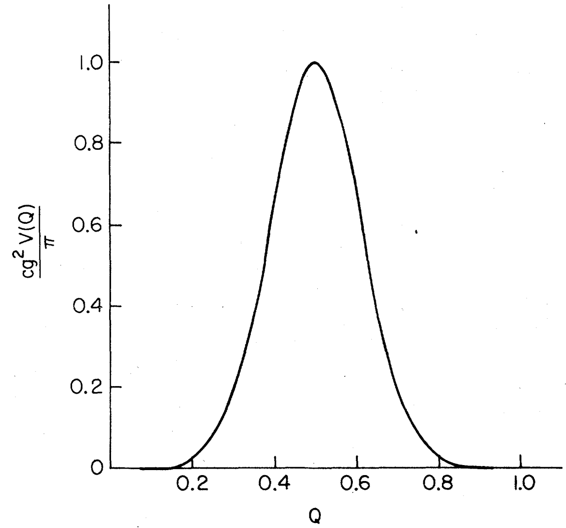


FIG. 6. The tunneling potential in winding number space.

$$\xi = \frac{\lambda^2}{\lambda^2 + c^2}. \tag{4.26}$$

The functions B_ξ and B are the incomplete and complete β functions, respectively. Clearly $Q(\lambda = -\infty) = 0$ and $Q(\lambda = \infty) = +1$. Equation (4.24) can be inverted numerically to express λ as a function of Q. The tunneling potential obtained from Eqs. (4.19) and (4.24) is shown as a function of Q in Fig. 6. Thus we can visualize the tunneling process as a particle of (Q-independent) mass M penetrating the potential barrier $V(Q)$ in winding number space. The tunneling rate is, of course, unchanged by the new parametrization.

C. Zero modes

In the next WKB approximation to the tunneling amplitude, the effects of so-called zero modes will be encountered. The next WKB approximation consists of calculating the factor which multiplies e^{-R_0} by including Gaussian fluctuations around an MPEP.[2] Zero modes arise when an MPEP is not isolated but is a member of a family of paths related by a symmetry operation. Fluctuations which correspond to moving from one MPEP to another in the same symmetry-related family is not damped, and the effects of such zero-mode fluctuations cannot be computed as Gaussian integrals. For example, an instanton centered at some other location than the origin in the x_1-x_2 plane corresponds to an MPEP in the same family as the MPEP associated with an instanton centered at the origin. The fluctuation which corresponds to moving the instanton infinitesimally is clearly not damped and gives rise to a zero mode. The treatment of these effects is now standard.[4] In this treatment, the integral over the coefficient of the zero mode is replaced by an integral over a collective coordinate, and a factor of group volume results in the tunneling amplitude. In the nonlinear σ model there are four such zero modes corresponding to the symmetry operations of space-time translation, scaling, and rotation around the 3 axis in spin space. Each of these operations will have a corresponding collective coordinate. In addition to the four symmetries just mentioned, there are two independent rotations which change the direction of the 3 axis in spin space. These operations do not change a collective coordinate associated with the instanton, but instead are a change in the boundary condition Eq. (3.3) requiring $\vec{s}$ to approach $(1/g)(1,0,0)$ at infinity. The divergence of the Gaussian fluctuation integral around an instanton which these rotations cause would also appear in computing the Gaussian fluctuations around the uniform background field $\vec{s} = (1/g)(1,0,0)$.[5] As pointed out by Jevicki,[5] if one normalizes the instanton amplitude by the vacuum persistence amplitude associated with the classical vacuum state $\vec{s} = (1/g)(1,0,0)$, the result will be finite. Of course physical quantities such as Green's functions are not affected by this divergence.

To summarize, we have seen that the instanton solutions can be used to calculate the tunneling amplitude between aligned spin configurations. The intermediate field configurations were found, the tunneling potential constructed, and the tunneling rate determined using the Minkowski-space formalism discussed earlier. Although we have given an explicit intermediate field configuration only for a winding number change of unity, a similar treatment obtains for the multi-instanton case.

V. MULTIPLE VACUUMS

In this section, we take up the question of the existence of multiple vacuums in the σ model. Let us start by reviewing the standard presentation of the situation in gauge theory.[6] In any gauge, the winding number q can be written as the integral of the winding number density ρ,

$$q = \frac{1}{32\pi^2}\int F^a_{\mu\nu}\,{}^*F^a_{\mu\nu}d^4x \equiv \int \rho\, d^4x\,, \tag{5.1}$$

where $F^a_{\mu\nu}$ is the usual field strength tensor and ${}^*F^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\rho}F^a_{\alpha\rho}$. Further, the density ρ can be expressed as the formal divergence of a current K_μ given by

$$K_\mu = \frac{1}{16\pi^2}\epsilon_{\mu\alpha\beta\gamma}(A^a_\alpha \partial_\beta A^a_\gamma - \tfrac{1}{3}\epsilon^{abc}A^a_\alpha A^b_\beta A^c_\gamma), \quad (5.2)$$

where K_μ satisfies

$$\partial_\mu K_\mu = \rho. \quad (5.3)$$

We use the term formal divergence as a reminder that there may be singular or discontinuous points in K_μ even when the density ρ is a smooth function.[7] However, there exist gauges, notably the temporal gauge, in which the time evolution is continuous, and no singularities or discontinuities of K_μ need occur. The additional assumption is usually made that the asymptotic behavior of the gauge fields is such that the contributions of surface at spatial infinity can be ignored. Under these conditions, Gauss's law gives

$$q = \lim_{x_4\to\infty}\int d^3x K_4(x_4, \vec{x}) - \lim_{x_4\to-\infty}\int d^3x K_4(x_4,\vec{x}), \quad (5.4)$$

so that winding number can be expressed as the difference of terms at large positive and negative times. This allows the winding number to be assigned to the classical vacuum configurations themselves and leads to the multiple- or θ-vacuum description of a gauge theory.

Now let us investigate whether there is any way to realize a multiple-vacuum picture in the nonlinear σ model. We begin by considering the way the value of q is built up for the $q=1$ pseudoparticle, applying Gauss's law to the space-time volume V between the surfaces $x_2=-T/2$ and $x_2=+T/2$. The limit $T\to\infty$, or equivalently $V\to\infty$, will be taken at the end. For the instanton, the current K_μ is not singular or discontinuous inside V so we have

$$q = \lim_{V\to\infty}\int_V \rho\, d^2x = \lim_{V\to\infty}\int_V \partial\cdot K\, d^2x = \lim_{V\to\infty}\int_\Sigma K\cdot d\Sigma, \quad (5.5)$$

where the surface Σ in the last equality includes the spatial surfaces normal to the 1 axis at $x_1 = \pm\infty$, as well as the constant time surfaces at $x_2=\pm T/2$. As $x\to\infty$ in any direction, the spin vector goes to the top of the unit sphere. Using this we may obtain q on Σ, which has $\theta\sim 0$, as

$$q = \lim_{T\to\infty}\left\{-\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\partial_1\Phi dx_1\right]_{x_2=-T/2}^{x_2=T/2} + \left[\frac{1}{2\pi}\int_{-T/2}^{+T/2}\partial_2\Phi\, dx_2\right]_{x_1=-\infty}^{x_1=+\infty}\right\}, \quad (5.6)$$

which evidently expresses q in terms of the total change in the aximuthal angle Φ in moving around the surface Σ. From the explicit formulas (3.16) and (3.17), which give θ and Φ for the $q=1$ instanton, we can easily see that the contribution on the $x_1=\pm\infty$ surfaces are negligible. This means the winding number comes solely from the large time surfaces, just as in the gauge theory case. On these constant time surfaces, the motion of the spin on the unit sphere traces out closed curves through the north pole as x_1 ranges from $-\infty$ to $+\infty$ (see Fig. 4). From the geometry of the curves or using the explicit functions θ and Φ from Eqs. (3.16) and (3.17) in Eq. (5.6) above, we can see that the value $q=1$ comes about as $\frac{1}{2}$ from the integral at $x_2=T/2$ and $-(-\frac{1}{2})$ from the integral at $x_2=-T/2$, in the limit $T\to\infty$. The case of a general value of q goes through in an entirely analogous way.

With these results in hand, it is tempting to try to set up a multiple-vacuum description by labeling the classical vacuum configurations with a number called "topological index" defined by

$$N = -\lim_{x\to\pm\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\partial_1\Phi(x_1,x_2)dx_1. \quad (5.7)$$

For the instanton solutions of winding number q, discussed above, denoting the initial and final values of N as N_i and N_f, the relationship

$$q = N_f - N_i \quad (5.8)$$

holds.

The topological index N defined by (5.7) and the winding number q are logically distinct quantities. The index is defined in terms of the curve at the top of the sphere which specifies the spin configuration at asymptotic early and late times. However, the winding number q is defined in terms of the solid angle covered on the sphere in the tunneling process defined over all space-time. Nevertheless, for the instanton solutions, the two are related by Eq. (5.8).

So far everything we have said is consistent with a multiple-vacuum interpretation. The critical questions in deciding whether a multiple-vacuum interpretation is really correct centers around Eq. (5.8), which so far holds only for pseudoparticle solutions. If, for example, $N_f = +\frac{1}{2}$ can *only* be reached from $N_i=-\frac{1}{2}$ by mapping the whole sphere once ($q=1$), then there is indeed a barrier between these two classical vacuums and N is a valid label for distinct vacuum configurations. The $q=1$ instanton would then only be distinguished among all $q=1$ mappings by giving the maximal tunneling rate. On the other hand, if $N_f=\frac{1}{2}$ could be reached from $N_i=-\frac{1}{2}$ without mapping the sphere at all ($q=0$), then vacuum configurations with $N=+\frac{1}{2}$ are not separated by a barrier and there would be no physical meaning to the label N. We now proceed to show that this latter possibility is the correct one.

Let us return to the situation described in Fig. 4. For $x_2=-T/2$, as x_1 ranges from $-\infty$ to ∞, the spin vector maps out a counterclockwise curve as shown with change in azimuthal angle Φ given by

$$\lim_{T\to\infty}\left(\Phi\left(\infty,-\frac{T}{2}\right)-\Phi\left(-\infty,-\frac{T}{2}\right)\right)=\pi, \tag{5.9}$$

corresponding to $N=-\frac{1}{2}$. Similarly, for $x=+T/2$, we have a clockwise curve with

$$\lim_{T\to\infty}\left(\Phi\left(\infty,\frac{T}{2}\right)-\Phi\left(-\infty,\frac{T}{2}\right)\right)=-\pi,$$

and $N=+\frac{1}{2}$. The $q=1$ instanton takes the early time curve into the late time curve by enlarging it, moving it down through the bottom of the sphere and up on the other side, finally arriving at the late time curve. However, we can transform the early time curve into the late time curve directly without ever leaving the top of the sphere. This is clear geometrically and we can express it analytically by parametrizing θ, Φ as follows:

$$\tan(\theta/2)=c/(x_1^2+T^2/4)^{1/2}, \tag{5.10}$$

$$\Phi=\left(\frac{2x_2}{T}\right)\tan^{-1}\left(\frac{T}{2x_1}\right). \tag{5.11}$$

We define Φ unambiguously by taking the argument of the inverse tangent to lie between 0 and π. These curves agree with those mapped out by the $q=1$ instanton for $x_2=\pm\frac{1}{2}T$, but for intermediate times, remain near the top of the sphere. The sequence of curves traced out as x_2 ranges from $-T/2$ to $+T/2$ is shown in Fig. 7.

Given the specific form of θ and Φ from Eqs. (5.10) and (5.11), it is straightforward to apply our previous formalism and calculate the tunneling amplitude. We omit the details: R turns out to be $O(1/T^3)$ and clearly vanishes as $T\to\infty$. Physically, this is obvious since in the limit $T\to\infty$ the spin vector remains vertically upward for all x_1. Therefore, a classical vacuum configuration of index $+\frac{1}{2}$ is not separated by any barrier from one of index $-\frac{1}{2}$. Similar arguments handle the case of other pairs of values of initial and final indices. Since the actual spin vectors are identical for classical vacuums of different index and as discussed above they are not separated by a potential barrier, it is clear that they are physically identical. In short, there is no evidence for multiple vacuums here, or put another way, the vacuum

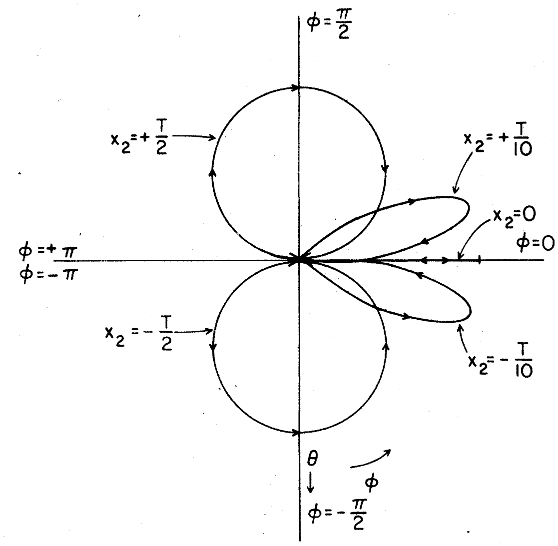


FIG. 7. The sequence of curves traced on the sphere for intermediate times by Eqs. (5.10) and (5.11).

is not a "Bloch wave" in the nonlinear σ model.

A final point of interest is to return to the formula for winding number q in terms of the current K_μ. Since as we have said the spin remains at the top of the sphere in the limit $T\to\infty$ of the process described by (5.10) and (5.11), the transition has $q=0$. From (5.6) we have

$$0=\lim_{T\to\infty}\left\{-\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\partial_1\Phi\,dx_1\right]_{x_2=-T/2}^{x_2=T/2}+\left[\frac{1}{2\pi}\int_{-T/2}^{T/2}\partial_2\Phi\,dx_2\right]_{x_1=-\infty}^{x_1=\infty}\right\}. \tag{5.12}$$

The first term contributes $+1$ here, just as for the $q=1$ instanton. Therefore it is clear that the spatial surface terms must be non-negligible in this case. Using (5.11) we have

$$\partial_2\phi=0 \text{ for } x_1=+\infty, \tag{5.13}$$

$$\partial_2\phi=\frac{2\pi}{T} \text{ for } x_1=-\infty, \tag{5.14}$$

and (5.12) is satisfied as

$$0=1+0-\frac{1}{2\pi}\left(\frac{2\pi}{T}\right)T. \tag{5.15}$$

This example shows that the winding number q will not always be expressible as the difference of two asymptotic time surface integrals. Spatial surface contributions may or may not be negligible. The cases where they are not negligible account for the failure of the relationship (5.8) between index and winding number to hold in general.

Note that no physical quantity (energy, momentum, etc.) is involved in these nonvanishing spatial surface terms.

To summarize, in the two-dimensional nonlinear σ model, the classical vacuum configuration with all spins up is a physically unique configuration, not a set of different configurations separated by potential barriers. As mentioned in the Introduction, the notion of tunneling is nevertheless still meaningful. A vacuum configuration may reach another vacuum configuration through a sequence in which the spin vector always remains near the top of the sphere, but the transition may also involve mapping the whole sphere one or more times. The latter processes are tunneling processes in the true sense of the word. They have the feature common in problems involving angular variables that the initial and final configurations are the same. These tunneling processes must be understood in order to develop a quantum theory of the σ-model ground state.[8]

The lack of a multiple or θ vacuum suggests a possible difference between the nonlinear σ model in $d=2$ and non-Abelian gauge theories in $d=4$, which are otherwise quite analogous in their properties. Before drawing this as a hard conclusion, however, it would seem worthwhile to reanalyze carefully the assumptions made in the usual treatment of gauge theory about the behavior of gauge fields at spatial infinity. We intend to explore this question in future work.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant No. NSF PHY77-25279.

---

*On leave from American University of Beirut, Lebanon.

[1]K. M. Bitar and S. J. Chang, Phys. Rev. D 17, 486 (1978).

[2]K. M. Bitar and S. J. Chang, Phys. Rev. D 18, 435 (1978).

[3]A. M. Polyakov, Zh. Eksp. Teor. Fiz. Pis'ma Red. 22, 503 (1975) [JETP Lett. 22, 245 (1975)].

[4]J. L. Gervais, A. Jevicki, and B. Sakita, Phys. Rep. 23C, 281 (1976).

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

[6]R. Jackiw, Rev. Mod. Phys. 49, 681 (1977). The original references are C. G. Callan, R. F. Dashen, and D. J. Gross, Phys. Lett. 63B, 334 (1976); R. Jackiw and C. Rebbi, Phys. Rev. Lett. 37, 172 (1976). Our $d=4$ indices are Euclidean and run from 1 to 4.

[7]See e.g., C. Bernard and E. Weinberg, Phys. Rev. D 15, 3656 (1977).

[8]The properties of the ground state of the nonlinear σ model are discussed in our companion paper. Report No. Ill-(TH)-78-17, 1978 (unpublished).