National Library of Canada — Bibliothèque nationale du Canada
Canadian Theses Division — Division des thèses canadiennes
Ottawa, Canada
K1A 0N4

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur
David Lorne Henty

Date of Birth — Date de naissance
25/10/1954

Country of Birth — Lieu de naissance
Canada

Permanent Address — Résidence fixe
62 Montcalm Ave.
Camrose, Alberta

Title of Thesis — Titre de la thèse
Topological Aspects of Non-Linear Field Theory

University — Université
University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée
M.Sc.

Year this degree conferred — Année d'obtention de ce grade
1979

Name of Supervisor — Nom du directeur de thèse
Dr. A. Z. Capri

Date
August 29/1979

Signature
David Henty

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

THE UNIVERSITY OF ALBERTA

TOPOLOGICAL ASPECTS OF NON-LINEAR FIELD THEORY

BY

© DAVID L. HENTY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF M.Sc.

IN

THEORETICAL PHYSICS

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled TOPOLOGICAL ASPECTS OF NON-LINEAR FIELD THEORY submitted by DAVID L. HENTY in partial fulfilment of the requirements for the degree of Master of Science.

....
Supervisor

....

....

....

Date August 27, 1979

To My Parents

ABSTRACT

This thesis reviews recent work in quantum field theory based on solutions to c-number field equations which possess topologically non-trivial boundary conditions. Chapter I introduces the basic concepts within the framework of several model field theories and outlines some relevant topological methods and results. In Chapter II the analysis is extended to pure Yang-Mills theory in Euclidean space and the pseudoparticle solution is derived. Chapter III introduces topological solitons and pseudoparticles into quantum field theory by utilizing the path integral formulation of quantum field theory. This chapter assumes no previous knowledge of path integral quantization and develops the formalism for both gauge fields and scalar and fermion fields. Chapter IV introduces possible applications of the analysis of the preceding chapters to elementary particle physics.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. A. Z. Capri, for suggesting this extremely interesting topic and also for encouragement and a great deal of patience. It is also a pleasure to acknowledge the benefit of enumerable discussions with fellow graduate students; in particular, Mr. Gordon Semenoff, Mr. Fred Salevsky, Mr. Pasquale Sodano, Mr. Eric Hiob and Mr. David Spaner. Finally, I would like to thank Mrs. Lee Cech for her excellent and speedy typing.

TABLE OF CONTENTS

CHAPTER		PAGE
I	Classical Field Configurations with Non-Trivial Topology	1
	1.1 Introduction and Motivation	1
	1.2 The ϕ^4 Kink	4
	1.3 The Vortex	14
	1.4 Topological Methods	24
	1.5 The Monopole or Hedgehog	34
II	Topological Configurations in Euclidean Space	47
	2.1 Introduction	47
	2.2 The Instanton or Pseudoparticle	51
	2.3 Multi-Instanton Configurations	60
	2.4 The Meron	63
	2.5 Discussion	65
III	Quantum Theory of Solitons	67
	3.1 Introduction	67
	3.2 Path Integral Quantization	68
	3.3 Incorporation of Gauge Fields	77
	3.4 Incorporation of Solitons	81
	3.5 Incorporation of the Instanton	91
	3.6 Fermion Zero Modes	98
IV	Applications	103
	4.1 Introduction	103
	4.2 The Vacuum in Non-Abelian Gauge Theories	103
	4.3 The U(1) Problem	109
	4.4 Mass Generation by Instantons	112

TABLE OF CONTENTS (Cont.)

CHAPTER	PAGE
4.5 Speculations on Quark Confinement	115
4.6 Conclusion	119
FOOTNOTES	123
REFERENCES	125

LIST OF TABLES

Page

Table 1. Homotopy Groups for Some Common Lie Groups 33

LIST OF FIGURES

FIGURE		PAGE
1	The field potential of the kink model as a function of ϕ.	6
2	The kink solution.	12
3	The field potential for the Abelian Higgs Model.	18
4	The form of the vortex solution.	23
5	The Higgs field superimposed on two dimensional space.	25
6	The gauge function as a map.	27
7	Homotopic maps on $E^2 - \{0,0\}$.	29
8	't Hooft's intuitive construction for a flux tube ending in a monopole.	38
9	The gauge transformation of the monopole into the Abelian gauge. The dotted lines in (a) indicate the direction of the rotation of the Higgs vectors required to reach the configuration in (b). These rotations clearly become undefined along the negative z axis.	45
10	Periodicity of the vacuum in non-abelian gauge theories.	106
11	Wilson's Quark Loop.	116

CHAPTER I

CLASSICAL FIELD CONFIGURATIONS WITH NON-TRIVIAL TOPOLOGY

1.1 Introduction and Motivation

Quantum field theory has traditionally been approached by solving the free field equations and then incorporating the interaction terms by a systematic expansion in powers of the coupling. This approach has proven extremely successful for quantum electrodynamics and that theory remains the most accurate theory known. The advent of non-Abelian gauge theories, however, has made it apparent that perturbation theory is inadequate to give a complete description of the corresponding quantum field theory. For example, the problem of quark confinement is expected to involve effects of gauge fields in regions where the coupling constant is extremely large or even infinite, and a perturbation expansion in the coupling becomes meaningless.

The inherent limitations of perturbation theory can be traced to the non-linear nature of quantum field theory. The perturbation expansion, in contrast, involves only the free fields which are solutions to the linear free field equations. Field configurations with properties peculiar to the non-linearity of the interacting field configurations may therefore be missed. It is just such configurations which may give the qualitatively new effects expected in non-Abelian gauge theories. This is especially true since the equations of motion for non-Abelian gauge fields without the presence of matter are in themselves non-linear, and asymptotically free fields don't exist.

Applied mathematicians have in fact been aware of qualitatively new effects in non-linear field theories for some time. The corresponding field configurations are commonly referred to as solitons or solitary waves and were first explored in hydrodynamics. These solitary waves can be found as exact solutions to certain non-linear field equations and have been shown to possess a number of peculiar properties[1]. In particular, solitary waves do not dissipate as do ordinary waves and if several isolated solitary waves are allowed to interact, they will emerge from the interaction region unchanged in form or velocity. The similarity of this behavior to that of elementary particles is certainly suggestive.

The complexity of the combined system of interacting field equations and commutation relations of a realistic quantum field theory makes an analagous search for exact solutions impractical. If the quantum fields are treated as c-numbers, however, exact solutions of the field equations can be found in some cases. These solutions exhibit many of the properties of solitary waves and are commonly referred to as solitons. These solitons, however, gain their stability in a completely different manner than the solitary waves of applied mathematics. Where the solitary waves, mentioned earlier, gain their stability from a balance between dispersion and self-interaction, the field configurations under consideration here are stable due to the peculiar topology of their boundary conditions.† These peculiar boundary conditions are made possible by the existence of degenerate minima for the field energy. Thus the field theories which exhibit these "topological solitons" are restricted to the classical versions of quantum field theories with spontaneously broken symmetry (although technically speaking, Chapter II deals with an exception to this rule).

Associated with their classical stability, topological solitons possess a conserved topological current and a corresponding topological charge. This charge is always forced to take on discrete values and is in many ways analagous to a conserved quantum number emerging at a purely classical level.

The present chapter explores some of the properties of these classical field configurations within the framework of a few model theories, emphasizing their topological properties. Chapter II also deals with solutions to c-number field equations which possess topologically non-trivial boundary conditions. These are considered in Euclidean space, however, and as will become clear later, they are, for this reason, not purely classical configurations when considered in Minkowski space.

Although such "classical" configurations are interesting in their own right, from the point of view of

---

†Relativistic field theories may also be constructed which exhibit soliton solutions which are stable by dynamical rather than topological mechanisms(2), but these will not be considered here.

mathematics, their primary significance arises from their role in the complete quantum theory, and this will be the concern of Chapters III and IV.

1.2 The ϕ^4 Kink

The simplest example of a classical field theory which exhibits topologically stable soliton type solutions is the ϕ^4 theory in 2 dimensions (1 space, 1 time). This can be described by the Lagrangian

$$L = \int dx\mathcal{L}(\phi) = \int dx[\tfrac{1}{2}\,\partial_\mu\phi\partial^\mu\phi - U(\phi)] \qquad (1.1)$$

with the field potential

$$U(\phi) = \frac{\lambda}{2}\,\phi^4 - m^2\phi^2 + \frac{m^4}{2\lambda} = \frac{\lambda}{2}\,(\phi^2 - \frac{m^2}{\lambda})^2$$

Here $\mu=0,1$ (t,x) and the integral is over 1 dimensional space.

This Lagrangian could be interpreted as describing the dynamics of scalar "mesons"† in a 1 dimensional world, however, the sign of the mass term is wrong. When this theory is used to illustrate spontaneous symmetry breaking a new "shifted" field ϕ' is defined

$$\phi' = \phi - (m^2/\lambda)^{\frac{1}{2}}$$

and the field potential is rewritten in terms of ϕ'

$$U(\phi') = \frac{\lambda}{2}\phi'^4 + 2m\sqrt{\lambda}\,\phi'^3 + \frac{(2m)^2}{2}\,\phi'^2$$

†Everything in this chapter is purely classical and any terminology borrowed from quantum field theory is used merely for convenience.

Thus the physical "meson" has mass 2m and a cubic self-interaction in addition to the original quartic self-interaction. Now, the original form of the Lagrangian was invariant under the transformation $\phi \to -\phi$ but the cubic self-interaction of the physical meson clearly breaks the symmetry $\phi' \to -\phi'$ and the properties of the physical mesons would conceal any evidence of the underlying reflection symmetry of the basic fields. That is, reflection symmetry has been spontaneously broken. In this instance, however, the objects of interest are not the mesons that would emerge from quantising the field ϕ or ϕ' but rather the clumps of field energy, solitons, which already look like particles in the classical theory. For this reason it is unimportant whether the Lagrangian is written in terms of ϕ or ϕ' and for convenience the original field ϕ will be used.

The classical field energy is given by

$$E = \int dx H = \int dx\left[\frac{\delta L}{\delta \dot{\phi}}\dot{\phi} - L\right]$$

$$= \int dx\left[\frac{1}{2}(\dot{\phi})^2 + \frac{1}{2}(\partial_1\phi)^2 + U(\phi)\right] \qquad (1.2)$$

where $\dot{\phi} = \partial_0\phi = \frac{\partial\phi}{\partial t}$ and $\partial_1\phi = \frac{\partial\phi}{\partial x}$.

The classical vacuum or ground state can be defined as the field configuration for which the energy is a minimum (redefining the field potential always makes possible putting this minimum at zero energy). From equation (1.2) it can be seen that a ground state must correspond to a field configuration with ϕ independent of space and time and equal to a zero of $U(\phi)$. Figure 1 shows the form of the field potential given in equation (1.1). The two minima at $\phi = \pm m/\sqrt{\lambda}$ correspond to degenerate ground states which break the $\phi \rightarrow -\phi$ symmetry of the Lagrangian. This degeneracy of the classical vacuum is a characteristic of all theories with spontaneously broken symmetry and is precisely what allows the existence of soliton type solutions.

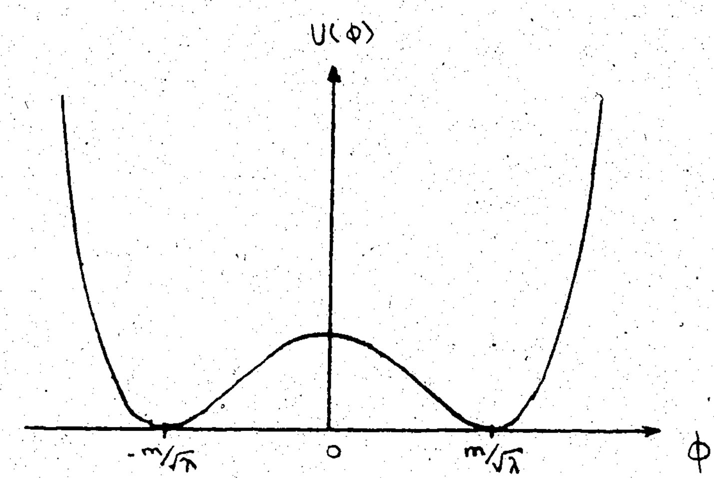


Figure 1. The field potential of the kink model as a function of ϕ.

Applying the Euler-Lagrange equations to the ϕ^4 Lagrangian gives:

$$\partial_\mu \partial^\mu \phi + \frac{\partial U}{\partial \phi} = (\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2})\phi + 2\lambda\phi^3 - 2m^2\phi = 0 \qquad (1.3)$$

Now, if the ϕ^4 topological solitons are to possess true soliton characteristics, they must represent stable, finite energy, non-dissipative solutions to the above equation. The last criterion can clearly be met if ϕ is restricted to be static, in which case (1.3) becomes

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial U}{\partial \phi} = 0 \qquad (1.4)$$

This can be integrated once to give:

$$\frac{1}{2}(\frac{\partial \phi}{\partial x})^2 = U(\phi) \qquad (1.5)$$

where the integration constant has been discarded to keep the energy integral finite. This allows the energy integral to be simplified to

$$E = \int dx[\frac{1}{2}(\partial_1\phi)^2 + U]$$

$$= \int dx(\partial_1\phi)^2 = \int dx[2U(\phi)]^{\frac{1}{2}} \qquad (1.6)$$

The requirement of (classical) stability can be made more precise by looking at the time dependent equation of motion (1.3). A time dependent solution to this equation can be written

$$\phi(x,t) = \phi_s(x) + \delta(x,t) \qquad (1.7)$$

where $\delta(x,t)$ is a time dependent perturbation added to the static solution $\phi_s(x)$. $\delta(x,t)$ can be expanded in a Fourier series:

$$\delta(x,t) = \sum_k e^{i\omega_k t} \psi_k(x)$$

Substituting (1.7) into (1.3) with the above form for $\delta(x,t)$ gives:

$$\Box\, \phi(x,t) + \frac{\partial U}{\partial \phi(x,t)} = 0$$

where to first order in $\delta(x,t)$, $\partial U/\partial\phi(x,t)$ is

$$\frac{\partial U}{\partial \phi(x,t)} = \left.\frac{\partial U}{\partial \phi}\right|_{\phi=\phi_s} + \left.\frac{\partial^2 U}{\partial \phi^2}\right|_{\phi=\phi_s} \delta(x,t) + \ldots$$

$$\therefore \Box\, \{\phi_s + \delta(x,t)\} + \frac{\partial U}{\partial \phi} + \frac{\partial^2 U}{\partial \phi^2}\, \delta(x,t) = 0$$

and $$\Box\, \{\sum_k e^{i\omega_k t} \psi_k(x)\} + \frac{\partial^2 U}{\partial \phi^2} \sum_k \psi_k\, e^{i\omega_k t} = 0$$

$$-\frac{d^2}{dx^2}\psi_k(x) + \left.\frac{\partial^2 U}{\partial \phi^2}\right|_{\phi=\phi_s} \psi_k(x) = +\omega_k^2\, \psi_k(x) \qquad (1.8)$$

Now this is just a one dimensional Schrodinger equation for a potential

$$\left.\frac{\partial^2 U}{\partial \phi^2}\right|_{\phi=\phi_s}$$

Thus the requirement of classical stability can be simply reduced to the demand that the Schrodinger equation (1.8) have non-negative eigenvalues. This assures that small perturbations will not grow in time. Equation (1.8) can also be derived by demanding that $\phi_s(x)$ be a local minimum of $E[\phi]$

in field space. That is

$$\left.\frac{\delta E[\phi]}{\delta\phi(x)}\right|_{\phi=\phi_s} = 0 = -\frac{d^2\phi_s}{dx^2} + \left.\frac{\partial U}{\partial\phi}\right|_{\phi=\phi_s}$$

is the requirement that $\phi_s(x)$ be a solution and thus the demand that

$$\frac{\delta^2 E[\phi]}{\delta\phi(x)\delta\phi(y)} = [-\frac{d^2}{dx^2} + \left.\frac{\partial^2 U}{\partial\phi^2}\right|_{\phi=\phi_s}]\,\delta(x-y) \qquad (1.9)$$

be a non-negative operator is the requirement that $\phi_s(x)$ be a local minimum. Studying the eigenvalue spectrum of this operator is again just the study of equation (1.8). The importance of classical stability as stated in this manner can be seen by noting that when the theory is finally quantized ϕ_s must represent a state which is not destroyed by quantum fluctuations.

An important property of the stability equation (1.8) is that it always possesses a zero eigenvalue associated with the so called translation mode. This can be seen by differentiating equation (1.4) with respect to x. This gives

$$[-\frac{d^2}{dx^2} + \left.\frac{d^2U}{d\phi^2}\right|_{\phi=\phi_s}]\,\phi_s' = 0 \qquad (1.10)$$

where $\phi_s' = \frac{d\phi_s}{dx}$. Thus ϕ_s' is a solution of equation (1.8) with $\omega=0$. The reason this solution is called the translation mode is that it follows from the translation symmetry of the theory. That is, if $\phi_s(x)$ is a solution so is $\phi_s(x+a)$ due to the arbitrariness of the choice of origin. But $\phi_s(x+a) = \phi_s(x) + a\phi_s'(x)$ to first order. Comparing this with equation (1.7) it is clear that $\phi_s'(x)$ represents a fluctuation with $\omega=0$. Similarly, any symmetry of the theory gives rise to a zero frequency mode since expanding ϕ_s along the symmetry direction in field space will give a $\delta\phi$ which represents a perturbation with $\omega=0$.

The above argument for the existence of the translation mode also holds for more than one spatial dimension. The stability equation (1.8), however, now possesses an n-fold degenerate translation mode corresponding to the freedom to make translations in any of the n directions. This conclusion allows a simple proof of an important theorem. The theorem states that for a theory with scalar fields obeying an equation of motion (1.3) (for arbitrary $U(\phi)$) there exist no nonsingular, time independent solutions of finite energy for space dimension greater than one. This theorem follows by merely recalling that for the usual Schrodinger equation the lowest eigenstate is nondegenerate. However for space dimension n greater than one, the zero frequency mode is n fold degenerate. Thus $\omega^2=0$ cannot be the lowest eigenvalue and solutions with $\omega^2<0$ must exist. Hence perturbations will grow in time and no static solutions can exist.

The above theorem was originally proven by Derrick[3] using a simple scaling argument. He demanded that static solutions be stable under transformations of the form $\phi(x) \rightarrow \phi(\lambda^{-1}x)$ which can be interpreted as a "stretching" of the soliton. For a static solution the energy of the soliton is given by

$$E = T + V$$

where $T = \frac{1}{2} \int (\vec{\nabla}\phi)^2 d^n x$ and $V = \int U(\phi) d^n x$

with n the number of space dimensions. Under the transformation $\phi(\vec{x}) \to \phi(\lambda^{-1}\vec{x})$

$$T \to \lambda^{n-2} T \quad \text{and} \quad V \to \lambda^n V$$

i.e. $$E(\lambda) = \lambda^{n-2} T + \lambda^n V$$

Now for stability, we require $\left.\frac{\partial E(\lambda)}{\partial \lambda}\right|_{\lambda=1} = 0$ which implies

$$(n-2)T + nV = 0$$

However, since both T and V are positive, this condition can only be satisfied for n=1. This shows that a static solution will minimize the energy integral only in one space dimension and hence in higher dimensions static "solutions" will not satisfy the field equations.

Returning to the simplified field equation for ϕ, equation (1.5), one can easily check that a space dependent solution is:

$$\phi_k = m/\sqrt{\lambda} \tanh[m(x-x_o)]$$

Inserting this in equation (1.6) gives

$$E = \frac{4}{3} m^3/\lambda$$

This shows a characteristic property of all soliton solutions, that is, the energy or mass of the soliton is always inversely proportional to the coupling constant. Figure 2 shows the form of the kink solution ϕ_k. The kink interpolates between different zeroes of $U(\phi)$ at $\pm\infty$ and its energy is concentrated near $\phi=0$; $x=x_o$. This figure also clearly shows the topological nature of the kink solution. Finite energy requirements fix $|\phi| \to m/\sqrt{\lambda}$ (zero of $U(\phi)$) as $x \to \pm\infty$. The discrete nature of the zeroes of $U(\phi)$ thus keeps a zero of ϕ with $E \neq 0$ trapped in the finite region of space. To allow the kink to dissipate would require continuously distorting ϕ over 1/2 of the 1-D space so that the value of ϕ_k corresponded in the $\pm$ directions. This would require $\frac{\partial\phi_k}{\partial t} \neq 0$ over an infinite region and hence infinite energy.

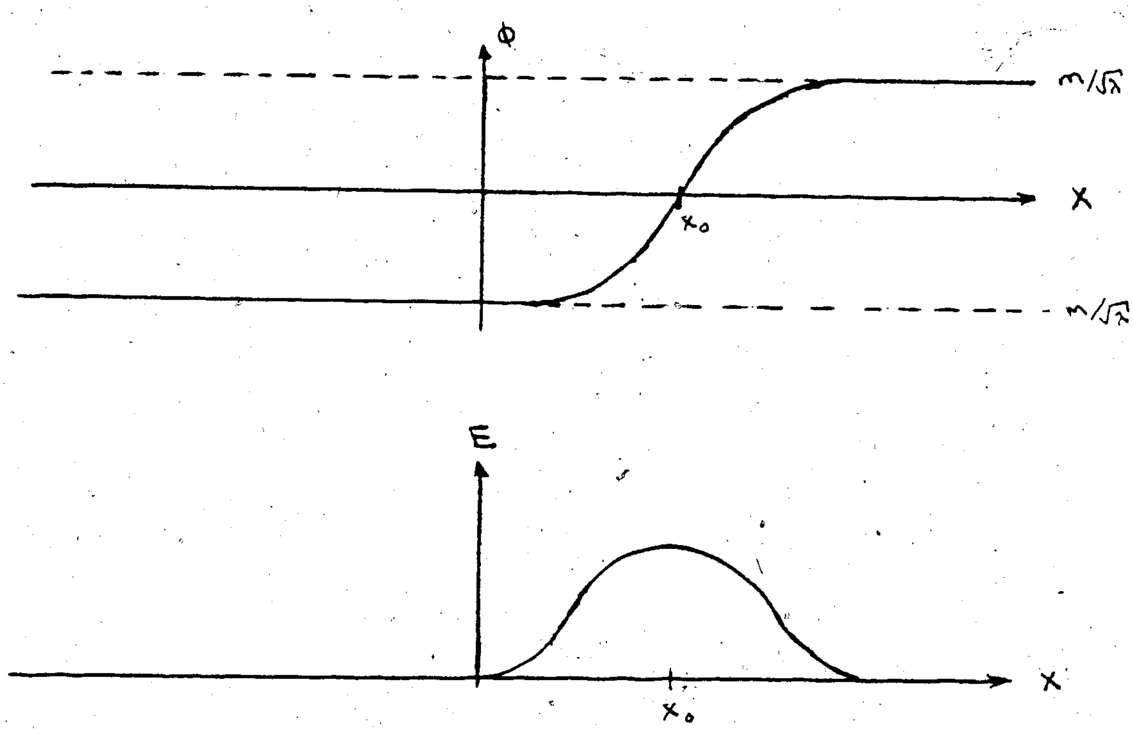


Figure 2. The kink solution.

It is possible to associate a conserved current J^{μ} with this topological conservation law

$$J^{\mu} = \varepsilon^{\mu\nu}\partial_{\nu}\phi$$

where $\varepsilon^{01} = -\varepsilon^{10} = 1$ and $\varepsilon^{00} = \varepsilon^{11} = 0$ and a corresponding topological charge or "quantum number"

$$k = \int dxJ^{0} = \int d\phi/dx\ dx = \phi(x)\Big|_{x\to+\infty} - \phi(x)\Big|_{x\to-\infty}$$

When the ϕ field is in the kink sector, ϕ takes different values as $x \to \pm\infty$ and k is non-zero, whereas when ϕ approaches the same "vacuum" as $x \to \pm\infty$, k=0 and there is no topological stability.

This intuitive understanding of the stability of the kink may be verified by substituting ϕ_k into the stability equation (1.8). This gives:

$$\left[-\frac{d^2}{dx^2} + 4m^2 - \frac{6m^2}{\cosh^2 mx}\right]\psi_k(x) = \omega_k^2\psi_k(x)$$

This equation can be solved exactly and is found to have a non-negative eigenvalue spectrum beginning with the zero frequency translation mode ω_o. Thus the kink solution is indeed stable.

The kink solution has been studied extensively, in both classical and quantum contexts, and a number of additional properties are known. Sources for a number of these results, as well as additional details on some of the above mentioned properties, may be found in Reference (4) to (8).

1.3 The Vortex

The kink solution found in the last section showed several very interesting properties, however, it is not clear how it can be generalized to more realistic theories in higher dimensions. Derrick's theorem clearly shows that searching for static solutions to scalar field equations in 2 or 3 space dimensions will have no success. Thus additional fields must be added and the logical choice would of course be to add gauge fields. The Lagrangian will therefore have the general form:

$$L = -\tfrac{1}{4} F^a_{\mu\nu} F^a_{\mu\nu} + (D_\mu\phi)^\dagger D^\mu\phi - U(\phi) \tag{1.11}$$

where $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f_{abc} A^b_\mu A^c_\nu$ (g=coupling constant)

where f_{abc} are the structure constants of the gauge group G, defined by:

$$[T^a, T^b] = i f^{abc} T^c$$

and where the T^a are the generators of the Lie group G. ϕ thus is a column matrix transforming under some representation of G. The covariant derivative D_μ is defined by (12):

$$D_\mu\phi = \partial_\mu\phi - igA^a_\mu T^a\phi$$

It will often be convenient to express the gauge fields in a form contracted over group indices:

$$A_\mu = A^a_\mu T^a \tag{1.12}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + g/i\ [A_\mu, A_\nu] \qquad (1.13a)$$

$$[D_\mu, D_\nu]\phi = igF_{\mu\nu}\phi \qquad (1.13b)$$

Also, it will often be convenient to work in the temporal gauge:

$$A_o^a = 0$$

In this gauge

$$D_o\phi = \partial_o\phi \quad \text{and} \quad F_{oi}^a = \partial_o A_i^a \quad i=1,2,3$$

The energy integral associated with (1.11) is:

$$E = \int d^n x\ \{\tfrac{1}{4} F^a_{\mu\nu}F^a_{\mu\nu} + (D_\mu\phi)^\dagger(D^\mu\phi) + U(\phi)\} \qquad (1.14)$$

The requirement that this integral remain finite, imposes the restriction

$$D_\mu\phi \xrightarrow[|\vec{x}|\to\infty]{} 0$$

i.e.

$$A_\mu(x) \xrightarrow[|\vec{x}|\to\infty]{} -\frac{i}{g}\ \phi^{-1}\partial_\mu\phi \qquad (1.15)$$

For the situations we will deal with, the fields $F_{\mu\nu}$ will be purely radial, hence $A_r(x) = 0$. Also we assume static fields $\partial_o\phi = 0$. ϕ can therefore have only an angular dependence as $|\vec{x}|\to\infty$. In 2-dimensions (1.15) becomes

$$A_\theta \xrightarrow[r\to\infty]{} -\frac{i}{g}\frac{1}{r}\ \phi^{-1}\ \frac{\partial\phi}{\partial\theta}$$

and in 3 dimensions the additional constraint

$$A_\psi \xrightarrow[r\to\infty]{} -\frac{i}{g}\ \phi^{-1}\ \frac{1}{r\sin\theta}\ \frac{\partial\phi}{\partial\psi}$$

with θ, ψ and r the standard spherical coordinates. Thus for large r, A_μ has in general a 1/r dependence.

Convergence of (1.14) also requires ϕ to go to a zero of $U(\phi)$ as $|\vec{x}|\to\infty$. If we denote a zero of $U(\phi)$ by $|\phi|=F$, ϕ can be written at large r as

$$\phi \to g(\Omega)\, F \tag{1.16}$$

where Ω denotes a general angular dependence in n dimensions and $g\varepsilon G$. Rewriting (1.15)

$$A_\mu(x) \to -\frac{i}{g}\, g(\Omega)^{-1}\partial_\mu g(\Omega) \tag{1.17}$$

we see that $A_\mu(x)$ goes to a pure gauge field at infinity and the vanishing of $F_{\mu\nu}$ is guaranteed. The finite energy requirements are therefore simply given by (1.16) and (1.17).

The simplest gauge theory of the form (1.11) which possesses topologically stable soliton solutions is the Abelian Higgs model in 2 space dimensions, first studied in this context by Nielsen and Olesen [(9)]. This model is just 2 dimensional scalar electrodynamics with the field potential of the Higgs scalars chosen so as to spontaneously break the Abelian U(1) gauge symmetry. In addition this model is formally a relativistic generalization of the Ginzburg-Landau equations of superconductivity with the Higgs scalars identified with the order parameter of superconductivity.[†] It is this correspondence which led Nielsen and Olesen to look for vortex solutions analagous to the

[†]For a formulation of superconductivity in terms of spontaneously broken phase invariance; see e.g. (10).

Abrikosov vortices or flux tubes of superconductivity. Although we are in 2-D we can turn our circular solitons into vortices by adding on a third dimension along which the fields are independent. Explicitly, the field potential is

$$U(\phi) = \frac{\lambda}{2}\,(\phi^*\phi - \mu^2/\lambda)^2 \tag{1.18}$$

where $$\phi = \phi_R + i\phi_I \quad .$$

$$\therefore L = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial_\mu + ieA_\mu)\phi^*(\partial^\mu - ieA^\mu)\phi + \mu^2\phi^*\phi - \frac{\lambda}{2}(\phi^*\phi)^2 - \mu^4/\lambda \tag{1.19}$$

The form of the potential $U(\phi)$ is shown in Figure 3. This clearly shows the U(1) symmetry corresponding to rotations about the origin. i.e.

$$\phi(x) \rightarrow \exp\{+i\chi(x)\}\phi(x)$$

and $$\phi^*(x) \rightarrow \exp\{-i\chi(x)\}\phi^*(x)$$

while $$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{e}\,\partial_\mu\chi(x)$$ since

we have allowed χ, the rotation parameter, to depend on x. $U(\phi)$ can be seen to have a form similar to $U(\phi)$ for the kink with however a continuous rather than discrete symmetry (quantized theory will have Goldstone bosons). Since the physical particles in the quantized theory will be associated with fluctuations around the zero energy configuration, we must define our classical vacuum as a point on the circle $\phi^*\phi = F^2 = \mu^2/\lambda$. To write the Lagrangian in terms of the "physical" fields, we reparametrize ϕ in terms of a "shifted"

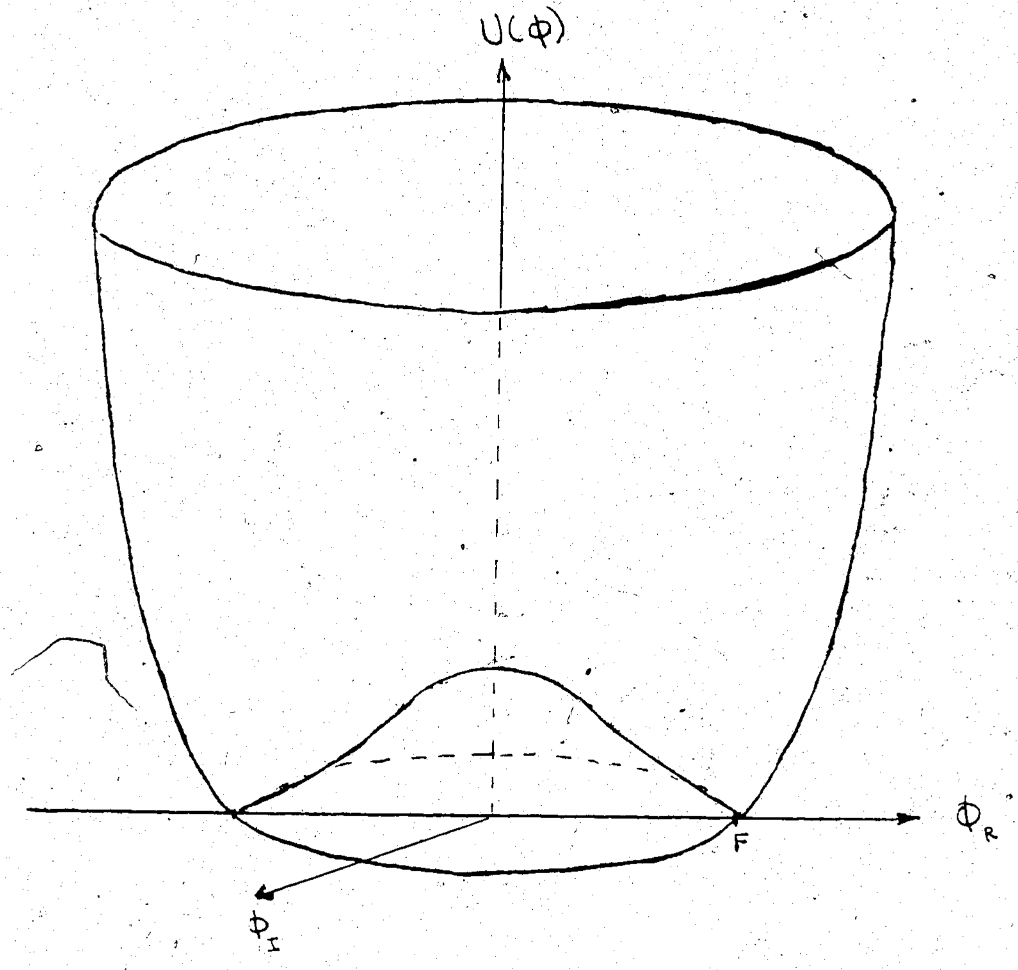


Figure 3. The field potential for the Abelian Higgs Model.

field η and a field ξ associated with the U(1) symmetry direction[12]:

$$\text{i.e.} \quad \phi = \exp\{i\xi/F\}(F+\eta)/\sqrt{2}$$

A U(1) rotation corresponds to a variation in ξ (i $\phi \to e^{-i\chi}\phi$ corresponds to $\xi \to \xi - F\chi$). It should noted that since ϕ is reparametrized in terms of 2 new fields, they are both real. Rewriting L in terms of ξ and η and keeping only first order terms (i.e. $\phi = 1/\sqrt{2}\,(F+\eta+i\xi+ \ldots)$) gives:

$$L = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \tfrac{1}{2}\partial_\mu\eta\partial^\mu\eta + \tfrac{1}{2}\partial_\mu\xi\partial^\mu\xi + \tfrac{1}{2}e^2F^2A_\mu A^\mu -$$
$$+ \sqrt{2}\,eFA_\mu\partial^\mu\xi - \mu^2\eta^2 + \ldots$$

ignoring higher order and constant terms. This shows that the η field can be interpreted (in the quantized theory) as a physical meson with mass $\sqrt{2}\mu$, however the would-be Goldstone boson mode ξ and the photon A_μ, have been mixed in an unusual way. To make the physical interpretation easier it is convenient to choose a gauge (usually called the unitary gauge or physical gauge) defined by

$$\phi' = e^{-i\xi/F}\phi$$
$$= (F + \eta)/\sqrt{2}$$

and

$$A'_\mu = A_\mu - \frac{1}{eF}\,\partial_\mu\xi$$

That is, the gauge has been chosen to absorb the ξ field and as a result the A_μ field has gained a longitudinal component

$\frac{1}{eF}\,\partial_\mu\xi$. Now since L is invariant under gauge transformations, L can be written directly in terms of ϕ' and A'_μ. Equation (1.19) becomes:

$$L = -\tfrac{1}{4}F'_{\mu\nu}F'^{\mu\nu} + (\partial_\mu + ieA'_\mu)\phi'^*(\partial^\mu - ieA^{\mu\prime})\phi' + \\ + \mu^2\phi'^*\phi' - \frac{\lambda}{2}\phi'^4 - \mu^4/\lambda$$

$$= -\tfrac{1}{4}F'_{\mu\nu}F^{\mu\nu\prime} + \tfrac{1}{2}\,\partial_\mu\eta\partial^\mu\eta + \tfrac{1}{2}F^2e^2A'_\mu A'^\mu \\ + \tfrac{1}{2}e^2A'^\nu_\mu\,(2F+\eta) - \tfrac{1}{2}\eta^2(3/2\lambda F^2 - \mu^2) \\ - \frac{\lambda}{2}F\eta^3 - \frac{\lambda}{8}\eta^4 \qquad (1.20)$$

Therefore, in this choice of gauge there appears an η meson with mass $= (3/2\lambda F^2 - \mu^2)^{\frac{1}{2}}$ and a massive vector meson with mass Fe and no massless Goldstone boson. The appearance of a mass for the photon implies the existence of a Meisner effect and hence the expectation that this model will possess solutions corresponding to tubes of flux confined by the Meisner effect as in superconductivity.

Since the vortex solutions we are looking for will be independent of the parametrization of our elementary fields, we can for convenience work with L in the form (1.19) rather than (1.20). The Euler-Lagrange equations:

$$\partial_\mu \frac{\partial L}{\partial(\partial_\mu\phi)} - \frac{\partial L}{\partial\phi} = 0$$

and

$$\partial_\mu \frac{\partial L}{\partial(\partial_\mu A^\nu)} - \frac{\partial L}{\partial A^\nu} = 0$$

imply the coupled field equations

$$(\partial_\mu + ieA_\mu)^2\phi = -2\lambda\phi^2\phi^* + 2\mu^2\phi \tag{1.21a}$$

and $$\partial^\nu F_{\mu\nu} = j_\mu = -\tfrac{1}{2}ie(\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*) + e^2A_\mu\phi^*\phi \tag{1.21b}$$

The latter implies

$$A_\mu = \frac{1}{e^2}\frac{j_\mu}{|\phi|^2} + \frac{1}{e}\partial_\mu\chi(x)$$

where ϕ has been taken to have the form

$$\phi = |\phi|\, e^{i\chi(x)}$$

The flux through a surface σ of our 2-D plane in the z direction is given by

$$\Phi = \int_\sigma F_{12}dxdy = \oint_\ell dx_i A^i \qquad i=1,2$$

Assuming $j_\mu = 0$ around the loop enclosing σ yields

$$\Phi = \frac{1}{e}\oint dx^i\partial_i\chi(x) = \frac{1}{e}\oint d\theta\,\frac{\partial\chi}{\partial\theta}$$

Single valuedness of ϕ requires $\chi(2\pi) = 2\pi n + \chi(0)$ and therefore

$$\Phi = \frac{1}{e}\{\chi(2\pi) - \chi(0)\} = \frac{2\pi n}{e}$$

This is the flux quantization familiar from type II superconductivity.

To find a soliton solution (vortex in 3-space) to this model would require finding an exact solution to the coupled equations (1.21a and b). Such a solution has never been found, however, in the general case.† The general form of the vortex can be found however, by choosing an appropriate Ansatz and looking at the equations for large r. Following Nielsen and Olesen, it is convenient to choose the gauge $A_o=0$; $\partial_i A^i=0$ and the cylindrically symmetric Ansatz (in 3-space, circular in 2-space)

$$A_i = \frac{\varepsilon_{ij} r_j}{|r|} |A(r)| \qquad i,j=1,2$$

i.e. $$\vec{A} = A(r)\, \hat{\theta}$$

and $$\phi = f(r)\, e^{i\theta}$$

Equations (1.21a and b) reduce to:

$$(a)\ -\frac{1}{r}\frac{d}{dr}\left(r\, \frac{df(r)}{dr}\right) + \left[\left(\frac{1}{r}-eA\right)^2 + 2\lambda\left(f(r)^2-2\mu^2/\lambda\right)\right]f(r)=0$$

$$(b)\ -\frac{d}{dr}\left(\frac{1}{r}\frac{d}{dr}(rA)\right) + f(r)^2\left(Ae^2 - \frac{e}{r}\right) = 0$$

If we now restrict ourselves to large r, equation (1.16) implies $f(r) \sim \text{const} = F$. From (b) one can now get the solution

$$A(r) \to \frac{1}{er} + C\, e^{-e\mu/\sqrt{\lambda}\, r} \qquad \text{large } r$$

i.e. $$\to \frac{1}{er} + C\, e^{-m_\nu r}$$

where m_ν is the photon mass. Thus the magnetic field of the vortex is non-neglible only over distances of the order of the penetration length $\delta = 1/m_\nu$. Also one can show that the

† An exact numerical solution has been found when an extra constraint on the coupling is imposed - reference (11).

distance over which the Higgs field ϕ varies appreciably from its vacuum value, is given by the coherence length $\xi=1/m_\eta$ where m_η is the mass of the η meson $\sqrt{2}\mu$.[9] The forms of A(r) and f(r) are shown in Figure 4.[9][11]

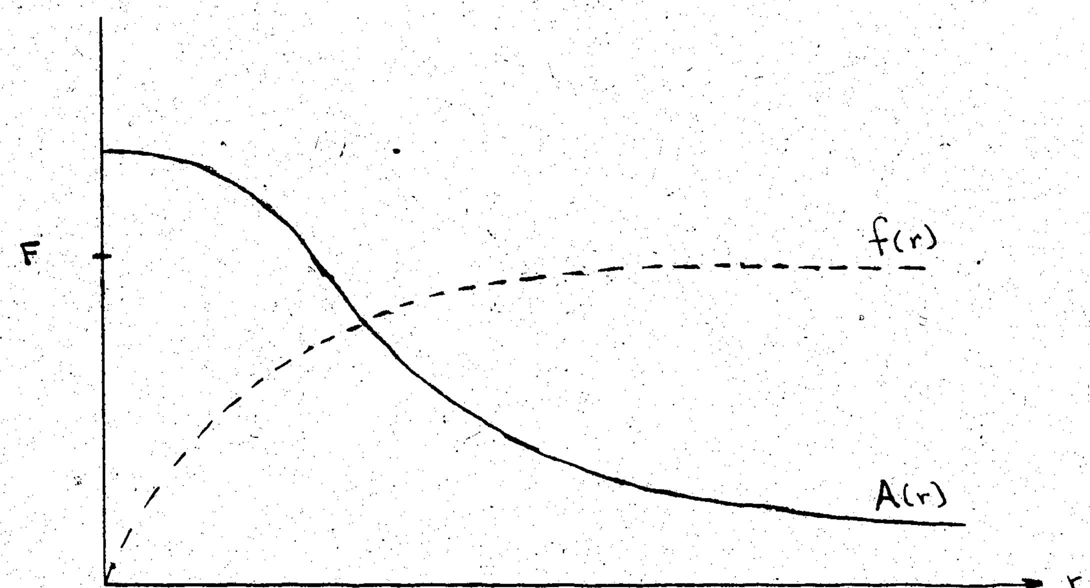


Figure 4. The form of the vortex solution.

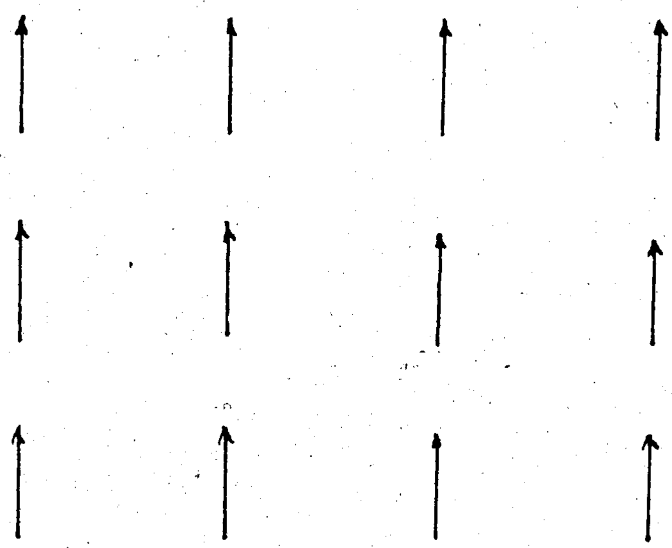
1.4 Topological Methods

The existence of field configurations with non-zero topological charge may be determined quite simply for any arbitrary field theory by utilizing methods from topology. Such an analysis does not in general prove the existence of solutions to the field equations, however. Nonetheless, a great deal of insight into the global properties of the theory may be gained and solutions, once found, may be classified by their topology.

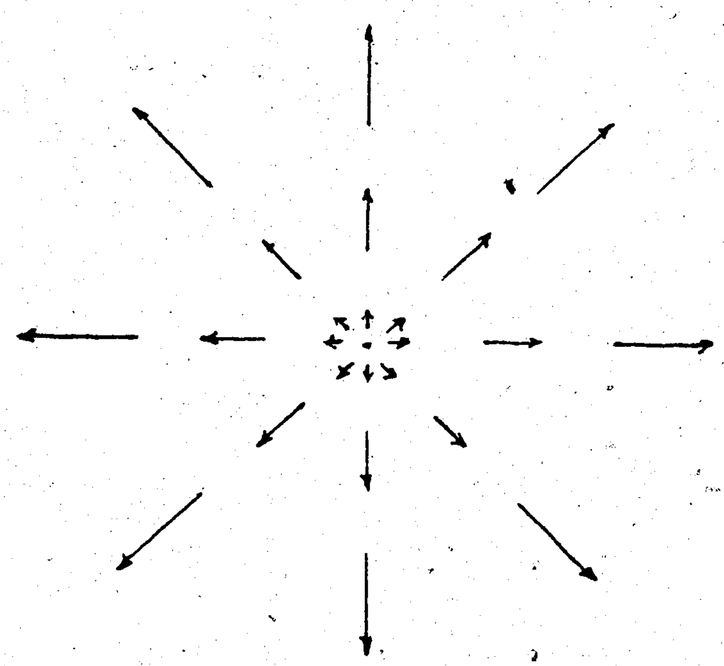
The role of such an analysis may be demonstrated by reconsidering the Abelian Higgs model of the last section. As pointed out above, the field ϕ must go to a zero of $U(\phi)$ for large r.

$$\phi \underset{r\to\infty}{\longrightarrow} F\, e^{i\chi(\theta)} \equiv Fg(\theta) \qquad (1.22)$$

Single valuedness of ϕ imposes the requirement $\chi(\theta) = n\theta$ or χ = constant. When χ equals a constant, the trivial solution ϕ=F over all space and A=0, satisfies the boundary condition (4.1) and the field equations (Figure 5(a)). Therefore any field configuration with this boundary condition can dissipate to the trivial, vacuum configuration. However, when χ= in one direction (i.e. θ=0) and -F in the other directic Continuity of ϕ implies that somewhere near the ori ace ϕ must have a zero. This is illustrated in Figu zero ϕ implies $U(\phi) \neq 0$, therefore the energy dens[illegible] of the field is non-zero over a region which can be chosen as centered at the origin. In other words,



(a) The vacuum sector with the Higgs vectors aligned over all space and $|\phi|=F$



(b) The one vortex sector where the nontrivial angular dependence of the Higgs field forces a zero at the origin. As in (a) the vectors are normalized to F at infinity.

Figure 5. The Higgs field superimposed on two dimensional space.

choosing a boundary condition for $\phi(x)$ or $g(x)$, with a non-trivial θ dependence implies the existence of a lump of energy at the origin.

The stability of this lump of energy is made evident by noting that it is impossible to continuously transform the boundary condition $\chi=\theta$ to χ = constant without breaking the requirement of single valuedness of ϕ. A similar argument for arbitrary n shows that possible field configurations fall into discrete sectors, depending on their boundary conditions and labelled by the integers n where

$$n = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{d\chi}{d\theta}\right)d\theta = -\frac{i}{2\pi}\int_0^{2\pi} g^{-1}(\theta)\left(\frac{dg(\theta)}{d\theta}\right)d\theta \;.$$

A field configuration with $n\neq 0$ cannot dissipate into the vacuum configuration with $n=0$ without developing a singularity (discontinuity) in its time development. n may therefore be considered an absolutely conserved topological charge. It is normally referred to as the winding number and is related to the flux of the vortex or lump via $\Phi = 2\pi n/e$.

The existence of nontrivial winding number for this model follows in a trivial way from a simple result in homotopy theory. The correspondence may be seen by noting that the gauge function $g(x)$ associates an element of the gauge group, in this case $U(1)$, for each value of x. That is, $g(x)$ is a mapping from space into the group manifold. In particular, $g(\theta)$ of equation (4.1) is a mapping from the boundary of 2-space onto a circle in the complex ϕ plane (Figure 6). In other words $g(\theta)$ is a mapping from one circle

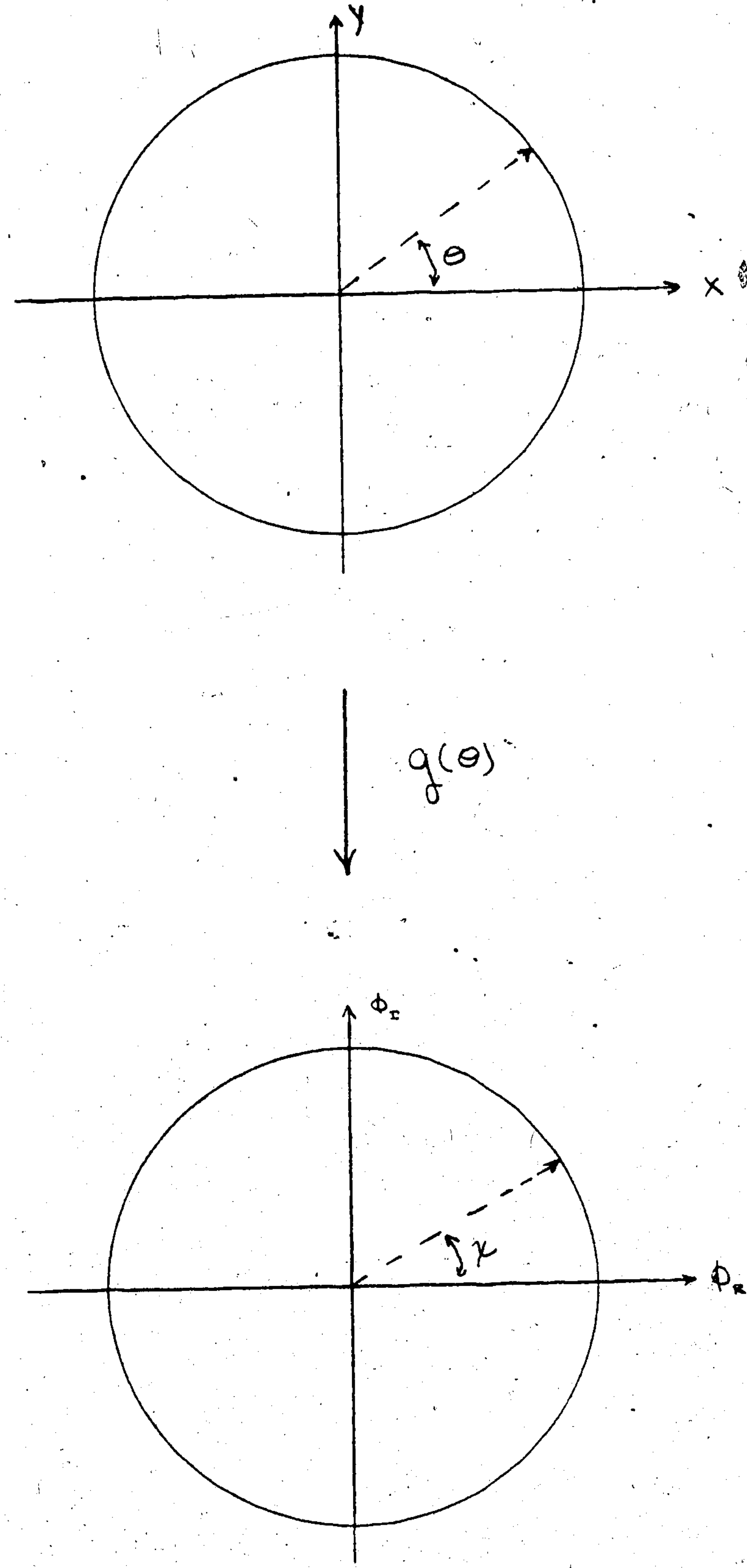
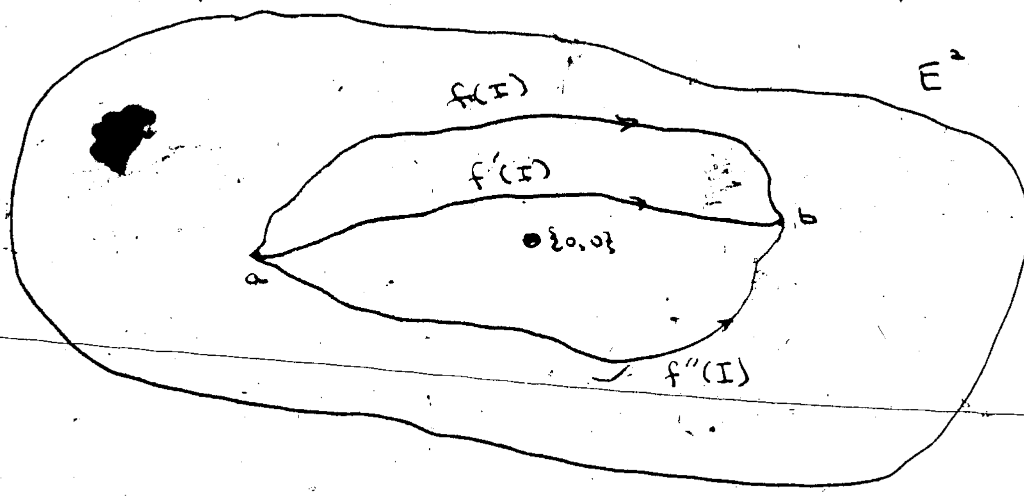


Figure 6. The gauge function as a map.

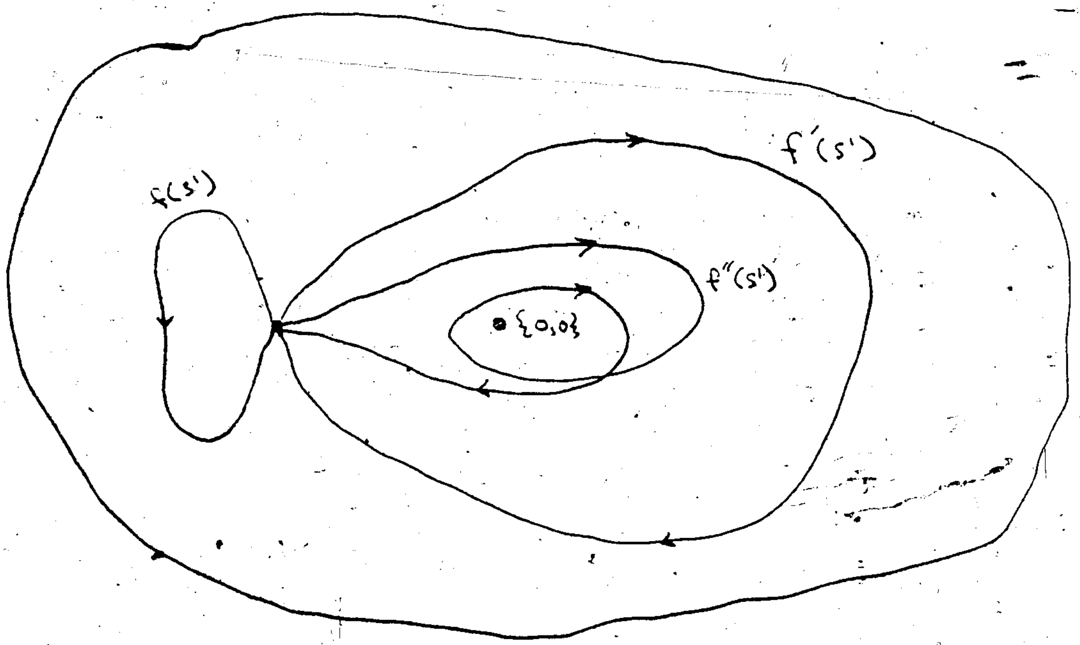
onto another, or in mathematical terminology, a mapping from S^1 to S^1. Such maps are known to fall into disconnected classes, called homotopy classes, labelled by the integers. Thus the existence of a topological charge in the Abelian Higgs model follows trivially from the mathematical result that maps from the boundary of 2-space (E^2) to the group manifold of U(1) fall into distinct homotopy classes.

The concept of homotopy class should be given a more careful treatment.† Two maps, f_0 and f_1, from a topological space X to a topological space Y ($f(x)=y$ for $x \epsilon X$, $y \epsilon Y$) are said to be homotopic, or members of the same homotopy class, if they can be continuously deformed into each other. More formally, f_0 and f_1, are homotopic if there exists a one parameter family of maps $f_t=F(x,t)$ such that $F(x,0)=f_0$ and $F(x,1)=f_1$ for all x. The family of maps f_t is called a homotopy from f_0 to f_1. A simple example is afforded by the maps f from I = [0,1], the unit interval of the real line, to $E^2 - \{0,0\}$; 2-space with the origin removed. The classification of these maps into homotopy classes may be understood in an intuitive manner by considering the images of the maps on $E^2 - \{0,0\}$. Two maps are then homotopic if the images of the unit interval can be continuously distorted into each other without breaking the path and keeping the endpoints fixed. Thus, in Figure 7(a) the maps f and f' with $f(0) = f'(0) = a$ and $f(1) = f'(1) = b$ are clearly homotopic

†Two standard mathematical references are (13); careful approaches for physicists may be found in (6), (14), (15) and (16).



(a)



(b)

Figure 7. Homotopic maps on $E^2 - \{0,0\}$.

whereas f" belongs to a separate homotopy class.

A more useful classification can be gained by identifying the ends of the unit interval. This leads to the classification of the maps of closed loops (topologically equivalent to S^1) onto the image space. In Figure 7(b), the loops f, f', and f" all belong to separate homotopy classes where the classes are distinguished by the number of times the loop encircles the origin.

A composition law · for maps may be defined. For example, in Figure 7(a)

$$f \cdot f'^{-1} = \begin{cases} f(2x) & 0 \leq x \leq \tfrac{1}{2} \\ f'^{-1}(2x-1) & \tfrac{1}{2} \leq x \leq 1 \end{cases}$$

$$\simeq e$$

Here e refers to the identity map, the map of the entire domain onto a single point and ≃ means "homotopic to". The inverse map is defined by

$$f^{-1}(x) = f(1-x)$$

and its image is merely the image of f(x) traced in the opposite direction.

With these definitions of composition and identity, the set of homotopy classes of maps of S^1, or the unit interval with endpoints identified, into a given topological space Y can be seen to have a group structure. This group is called the fundamental group of Y and is denoted $\Pi_1(Y)$.

Thus, for the example of Figure 7(b), Y is $E^2 - \{0,0\}$ and $\Pi_1(Y) = Z$, the additive group of integers.

In an exactly analagous manner, maps of S^2, or I^2 with boundary points identified, onto a given space Y can be placed in homotopy classes and form a group $\Pi_2(Y)$. Similarly maps of S^n onto Y form a group $\Pi_n(Y)$. It should be noted however that if the image space Y is not path connected, the homotopy groups must be defined at a given point in Y, and are denoted $\Pi_n(Y,x_o)$ (13).

Returning to the two dimensional Abelian Higgs model, it can now be seen that the discrete sectors of solutions are simply labelled by the elements of $\Pi_1(U(1)) = \Pi_1(S^1) = Z$. The elements of Z are just the winding numbers which distinguish homotopically inequivalent maps $g(\theta)$.

This result generalizes immediately to any gauge group. If G is the gauge group and H the subgroup which leaves the classical vacuum invariant, then, ignoring the possibility of other internal symmetries, the factor group G/H acts transitively on the space of zeros of $U(\phi)$. That is, if ϕ_o is a zero of $U(\phi)$, then any other zero may be given by $\tilde{g}\,\phi_o$ for $\tilde{g} \varepsilon G/H$. Thus, in a space of dimension d, the requirement of finite energy (or action in 4-dimensions) imposes the restriction

$$\phi \underset{r\to\infty}{\longrightarrow} \tilde{g}(\Omega)\,|\phi_o| \qquad \tilde{g}\ \varepsilon\ G/H$$

where Ω denotes all angular variables. The homotopy classes of the mapping $\tilde{g}(\Omega)$ then divide the space of nonsingular field configurations ϕ into discrete sectors labelled by $\Pi_{d-1}(G/H)$.

The existence of topological conservation laws may thus be deduced for any gauge field theory in any number of dimensions, merely by inspecting the appropriate $\Pi_n(Y)$. If the homotopy group has only one element $\{e\}$, then the field theory possesses only one sector, the vacuum sector, and any lumps of field energy may dissipate into the vacuum. If $\Pi_{d-1}(G/H)$ has more than one element, however, the field theory possesses a corresponding number of discrete sectors and field configurations in sectors other than the vacuum sector will not be able to dissipate.

Mathematicians have classified the $\Pi_n(Y)$ for Y the group manifold of a large number of Lie groups. Most of the results likely to be useful in physics are summarized in Table 1.(7),(13),(19) Some additional results are(14),(15),(17)

$$\Pi_1(SU(N)/Z_n) = Z_n \tag{1.23a}$$

$$\Pi_2(G/H) = \Pi_1(H) \text{ for G any simply connected, compact Lie group} \tag{1.23b}$$

$$\Pi_n(S^n) = Z \tag{1.23c}$$

$$\Pi_{n+1}(S^n) = Z_2;\ n \geq 3 \tag{1.23d}$$

$$\Pi_{n+2}(S^n) = Z_2;\ n \geq 3 \tag{1.23e}$$

$$\Pi_n(S^3) = \Pi_n(S^2);\ n > 2 \tag{1.23f}$$

$$\Pi_n(A \times B) = \Pi_n(A) \times \Pi_n(B) \tag{1.23g}$$

TABLE 1. Homotopy Groups for Some Common Lie Groups

Y	U(1)	U(2)	N≥3 U(N)	SU(2)	N≥3 SU(N)	SO(3)	SO(4)	SO(5)	SO(6)	N≥7 SO(N)	SP(N)
Π_1	Z	Z	Z	0	0	Z_2	Z_2	Z_2	Z_2	Z_2	0
Π_2	0	0	0	0	0	0	0	0	0	0	0
Π_3	0	Z	Z	Z	Z	Z	Z+Z	Z	Z	Z	Z
Π_4	0	Z_2	0	Z_2	0	Z_2	Z_2+Z_2	Z_2	0	0	Z_2

Also, denoting the simply connected covering group of G by $\tilde{G}$ and defining a subgroup C of the center of $\tilde{G}$ by $G = \tilde{G}/C$, then

$$\Pi_1(G/H) = C \quad \text{and} \quad \Pi_1(\tilde{G}/H) = 0 \qquad (1.23h)$$

1.5 The Monopole or Hedgehog

The topological analysis of the preceding section indicates that for a gauge field theory in three spatial dimensions, the criterion for the existence of topological conservation laws is that $\Pi_2(G/H)$ be non-trivial. For the unified models of the weak and electromagnetic interactions, or the grand unified models which incorporate strong interactions (e.g. SU(5)[19]), the gauge group G always contains an unbroken U(1) subgroup. Hence, for these models U(1) may be identified with H (or a subgroup of H for the grand unified models). If G has a compact covering group then the relation (1.23b)

$$\Pi_2(G/H) = \Pi_1(H)$$

leads to a consideration of $\Pi_1(U(1))$. Since this has non-trivial elements labelled by the integers, topologically stable field configurations should exist.

Unfortunately, the most popular model of weak and electromagnetic interactions, the Weinberg-Salam model, is based on the gauge group $SU(2)\otimes U(1)$ which does not have a compact covering group. Relation (1.23b) does not apply and instead

$$\Pi_2(SU(2)\otimes U(1)/U(1)) \equiv \Pi_2(SU(2)) = \{e\}$$

and no solitons are expected.

It is very possible though, that observed phenomenology is due to the complicated breaking of some larger compact group for which (1.23b) holds. For such a group broken to U(1), the expected solitons have indeed been found, first by 't Hooft[20] and Polyakov[21] independently for G = SO(3), and later by a number of others for more general groups[22]-[25] (reference (24) contains extensive references). These solitons have been shown to be stable magnetic monopoles associated with the unbroken U(1) electrodynamics. In this sense they are essentially Dirac monopoles with however, the added feature of finite energy. In addition, the unphysical Dirac string is not required in all gauges.

The simplest model in which to exhibit the properties of these monopoles is the gauge group SO(3) considered as in the treatment of 't Hooft. The gauge group SO(3) was originally proposed by Georgi and Glashow [26] as the basis of a model to unify weak and electromagnetic interactions without the introduction of weak neutral currents. It was abandoned when weak neutral currents were discovered but remains a useful model since it contains all the elements of unified theories based on larger compact groups. Along with the pure SO(3) gauge fields, the model contains an isotriplet of Higgs scalars which break the gauge symmetry and fermions which may be excluded for this analysis. The Lagrangian density is given by

$$L(x) = -\tfrac{1}{4}\, G^a_{\mu\nu} G^{a\mu\nu} + \tfrac{1}{2}\, (D_\mu \phi^a)(D^\mu \phi^a) - \frac{\lambda}{4}\, (\phi^a \phi^a - \mu^2/\lambda)^2 \qquad (1.24)$$

$$a = 1,2,3$$
$$\mu,\nu = 0,1,2,3$$

where $$D_\mu \phi^a = \partial_\mu \phi^a + e\, \varepsilon^{abc}\, A^b_\mu \phi^c$$

and $$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e\, \varepsilon^{abc}\, A^b_\mu A^c_\nu$$

The degenerate "vacua" of this theory are defined by the sphere in isospace satisfying $\phi^a \phi^a = \mu^2/\lambda \equiv F^2$. If a unique "vacuum" is chosen by taking the Higgs vector as pointing in a specific direction (e.g. the z direction), then the group of SO(2) rotations in the plane perpendicular to this direction, defines the little group H. That is, a SO(2) $\equiv$ U(1) symmetry remains unbroken with a corresponding massless vector particle (photon) A^3_μ. The other two gauge bosons become massive via the Higgs mechanism.

Now G/H = SO(3)/SO(2) is topologically equivalent to S^2 and therefore $\Pi_2(G/H) = \Pi_2(S^2) = Z$. Equivalently, $\Pi_2(G/H) = \Pi_1(H) = \Pi_1(SO(2)) = Z$. This theory consequently possesses an infinite number of topologically distinct sectors which can be classified by the wrapping number associated with the map $S^2 \to S^2$. The fields in these sectors have been shown to correspond to magnetic monopoles associated with the unbroken U(1) gauge symmetry. An intuitive argument showing why magnetic monopoles should appear when electrodynamics is embedded in a non-Abelian group, has been given by 't Hooft[20].

Consider a tube of magnetic flux entering a spherical region as shown in Figure 8. Outside the tube $F_{\mu\nu} = 0$ but there must exist a vector potential $\vec{A}$ such that

$$\int_{c_o} \vec{A} \cdot d\vec{x} = \Phi = \text{flux} = 2\pi n/e$$

Here electrodynamics has been considered embedded in the non-Abelian group (but $F_{\mu\nu} \neq G_{\mu\nu}$; see below) and the gauge condition $A_o = 0$ has been adopted. In the region outside the flux tube, $\vec{A}$ is given by

$$\vec{A}_c = \frac{i}{e}\, g(\theta)^{-1} \vec{\nabla} g(\theta) = \frac{1}{e}\vec{\nabla}\, \chi(\theta)$$

where the subscript c indicates that $\vec{A}$ can depend on the choice of curve c. As before $g(\theta)_c$ may be considered a map onto G/H from the curves c_o, c, ... If $g(\theta)_c$ may be varied continuously, as the curves c are contracted to a point at the bottom of the sphere, then the flux tube may end inside the sphere. This leads to a consideration of $\Pi_1(G/H)$. Since $SO(3) \equiv SU(2)/Z_2$, $\Pi_1(SO(3)/U(1)) = Z_2$ and flux tubes with

$$\Phi = 2\pi n/3 \qquad n = 2,4,6 \ldots$$

may terminate with magnetic monopoles at their endpoints. Since flux equals 4Π × magnetic charge, these monopoles carry magnetic charge $g_m = 1/e$. For the case of SU(2) broken to U(1), $\Pi_1(G/H) = \{e\}$ and monopoles carry charge $g_m = \frac{1}{2e}$, which is Dirac's condition.

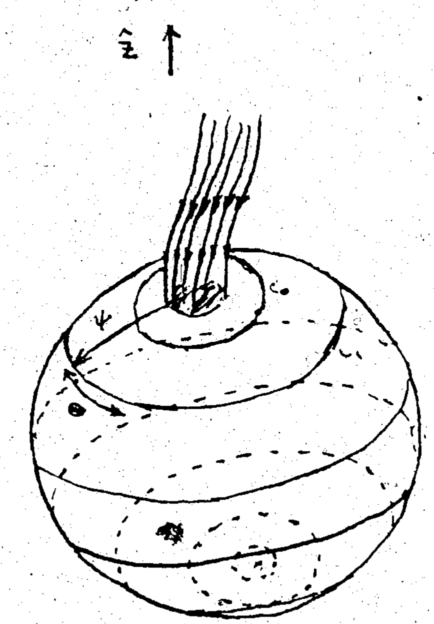


Figure 8. 't Hooft's intuitive construction for a flux tube ending in a monopole.

The possibility of the existence of stringless monopoles in non-abelian gauge theories may also be seen from a simple comparison with normal electrodynamics. In normal abelian electrodynamics, the requirement that $\vec{B}$ can be written in terms of a vector potential via

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

implies the condition

$$\vec{\nabla} \cdot \vec{B} = 0$$

Magnetic monopoles would spoil this however since

$$\vec{\nabla} \cdot \vec{B} = 4\pi \rho_m$$

where ρ_m is the monopole charge density. A Dirac string must therefore be introduced so that

$$\vec{\nabla} \cdot \vec{B}' = 0$$

where $$\vec{B}' = \vec{B}_{monopole} + \vec{B}_{string}$$

For non-abelian theories

$$\vec{B}^a = \vec{\nabla} \times \vec{A}^a + g f^{abc} \vec{A}^b \times \vec{A}^c$$

and the above analysis does not apply. Introduction of a magnetic charge does therefore not require abandoning the vector potential or introducing an unphysical string.

Returning to the problem of finding solutions for the Georgi-Glashow model; the field equations which follow from the Lagrangian (1.24) are

$$D^{\mu}G^{a}_{\mu\nu} = - e\ \varepsilon^{abc}\phi^{b}D_{\nu}\phi^{c} \tag{1.25a}$$

and

$$D^{\mu}D_{\mu}\phi^{a} = - \lambda\phi^{a}(\phi^{b}\phi^{b} - F^{2}) \tag{1.25b}$$

Making the spherically symmetric Ansatz

$$\phi^{a} = x^{a}H(r)/er^{2}$$

$$A^{a}_{o} = 0 \qquad A^{a}_{i} = \varepsilon_{aij}x^{j}[1-K(r)]/er^{2}$$

The field equations (1.25) now become

$$r^{2}\ \frac{d^{2}K(r)}{dr^{2}} = K(r)(K^{2}(r)-1) + K(r)H^{2}(r) \tag{1.26a}$$

and

$$r^{2}\ \frac{d^{2}H(r)}{dr^{2}} = 2H(r)K^{2}(r) + \frac{\lambda}{e^{2}}(H^{2}(r) - F^{2}e^{2}r^{2})H(r) \tag{1.26b}$$

Although it is known that these equations possess exact solutions of finite energy(28), the only exact analytic solutions that have been obtained were found for the limit $\lambda\to 0$(27). The corresponding K and H are

$$K(r) = eFr/\sinh(eFr)$$

$$H(r) = eFr\ \coth(eFr) - 1$$

Although this solution is essentially trivial, it retains enough of the complete theory to provide a testing ground for asymptotic and numerical investigations of the complete equations.

't Hooft and Polyakov considered the asymptotic form of the solutions to the complete equations. Finite energy requirements impose the conditions

$$|\phi| = F \quad \text{as} \quad r \to \infty$$

and therefore

$$H(r) \xrightarrow[r\to\infty]{} eF\, r$$

Also ϕ must be covariantly constant

$$D_i\phi \xrightarrow[r\to\infty]{} 0$$

which implies

$$A_i^a \xrightarrow[r\to\infty]{} \frac{c}{r} \qquad c=\text{constant}$$

Substitution of these asymptotic forms into the equations of motion leads to

$$K(r) \xrightarrow[r\to\infty]{} 0$$

and thus at large r, the Higgs and vector fields are

$$\phi^a = F\frac{x^a}{r} \qquad (1.27a)$$

and

$$A_i^a = \varepsilon_{aib}\frac{X_b}{er^2} \qquad (1.27b)$$

To exhibit clearly the physical nature of these fields it is necessary to clarify the connection with normal U(1) electrodynamics. Since this SO(3) theory possesses an invariant U(1) gauge group, there must exist a field tensor formally equivalent to that of Abelian electrodynamics, which can be written in an SO(3) invariant form. This tensor was given by 't Hooft[20]:

$$F_{\mu\nu} = \hat{\phi}^a G^a_{\mu\nu} - \frac{1}{e}\,\varepsilon_{abc}\,\hat{\phi}^a D_\mu \hat{\phi}^b D_\nu \hat{\phi}^c \tag{1.28}$$

where $$\hat{\phi}^a = \phi^a/(\phi^b\phi^b)^{\frac{1}{2}}$$

This clearly shows that the choice of orientation of the Higgs vectors, determines the structure of the Abelian subgroup.

Substitution of the asymptotic fields (1.27) into $F_{\mu\nu}$ gives

$$F_{ij} = \varepsilon_{ijk}\, X^k/er^3$$

and $$B_i = \tfrac{1}{2}\varepsilon_{ijk}\, F_{jk} = x_i/er^3$$

which corresponds to the field of a magnetic monopole with magnetic charge $g_m = 1/e$.

A magnetic current K_μ may be associated with this charge by the relation(29)

$$K_\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma}$$

Now, $F_{\mu\nu}$ may be rewritten in the form

$$F_{\mu\nu} = M_{\mu\nu} + H_{\mu\nu}$$

with $M_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ where $B_\mu = \hat{\phi}^a\, A^a_\mu$

and $$H_{\mu\nu} = \frac{1}{e}\,\varepsilon_{abc}\hat{\phi}^a\partial_\mu\hat{\phi}^b\partial_\nu\hat{\phi}^c$$

K_μ then becomes

$$K_\mu = \tfrac{1}{2}\,\varepsilon_{\mu\nu\rho\sigma}\partial^\nu H^{\rho\sigma}$$
$$= \frac{1}{2e}\,\varepsilon_{\mu\nu\rho\sigma}\varepsilon_{abc}\partial^\nu(\hat{\phi}^a\partial^\rho\hat{\phi}^b\partial^\sigma\hat{\phi}^c)$$

since $\varepsilon_{\mu\nu\rho\sigma}\partial_\mu M_{\rho\sigma} = 0$ if there are no singularity lines in the gauge field (that is, if $[\partial_\mu, \partial_\nu]B_\rho = 0$). Now the divergence of K_μ vanishes

$$\partial_\nu K^\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\partial^\mu\partial^\nu H^{\rho\sigma}$$
$$= 0$$

again assuming ∂_μ and ∂_ν commute, and the current K_μ is conserved. Thus the magnetic charge

$$g_m = \frac{1}{4\pi}\int k_0 \, d^3x \tag{1.29}$$
$$= \frac{1}{8\pi e}\int \varepsilon_{abc}\varepsilon_{ijk}\partial_i(\hat{\phi}^a\partial_j\hat{\phi}^b\partial_k\hat{\phi}^c)d^3x$$

satisfies $\dot{g}_m = 0$

Inserting ϕ^a given by equation (1.27a) into the expression for g_m gives $g_m = \frac{1}{e}$ as expected.

The dependence of K_μ solely on the Higgs vectors is somewhat surprising but serves to emphasize the topological rather than dynamical nature of the magnetic current. It is the peculiar alignment of the Higgs vectors as pointing radially outward, when isospace is superimposed on real space, that gives the monopole its stability and it is this property of the Higgs field which prompted Polyakov to dub this solution the hedgehog. The magnetic charge as given by equation (1.29) just classifies the map of the sphere defined by the Higgs vectors onto the sphere bounding 3-space, by the wrapping number (which is isomorphic to the elements of $\Pi_2(S^2)$)[29].

As was mentioned above, the vanishing of the vector field contribution to the magnetic current only occurs when the field has no singularities. In fact, the magnetic charge may be "transferred" completely from the Higgs field to the vector field by a "singular" gauge transformation $g(x)$. The necessity of a line along which $g(x)$ is singular is intuitively illustrated in Figure 9.

For the asymptotic form of the Higgs and vector fields, applying the gauge transformation

$$g(x) = \exp(-i\theta T_3)\exp(i\psi T_2)\exp(i\theta T_3)$$

where $T_a = \sigma_{a/2}$ and $\theta = \arctan(x_2/x_1)$ and $\psi = \arccos(x_3/r)$ yields the transformed fields

$$\hat{\phi}' = (\hat{\phi}^a T^a)' = g(x)(\hat{\phi}^a T^a)g(x)^{-1} = \delta_{a3} T_a$$

$$A_i' = (A_i^a T^a)' = g(x) A_i g^{-1}(x) + \frac{i}{e}\, g(x)\partial_i g^{-1}(x)$$

$$= \delta_{a3}\, \frac{1}{e}\, \varepsilon_{i3k}[x^k/r(r-x^3)]T_a$$

In this gauge, the Higgs vectors are all aligned in one direction and only one isospin component of the vector field survives. Also, the electromagnetic field tensor is now

$$F_{\mu\nu} = \partial_\mu A_\nu^3 - \partial_\nu A_\mu^3$$

The fields are all, therefore, formally equivalent to normal electrodynamics, and this gauge is correspondingly referred to as the Abelian gauge.

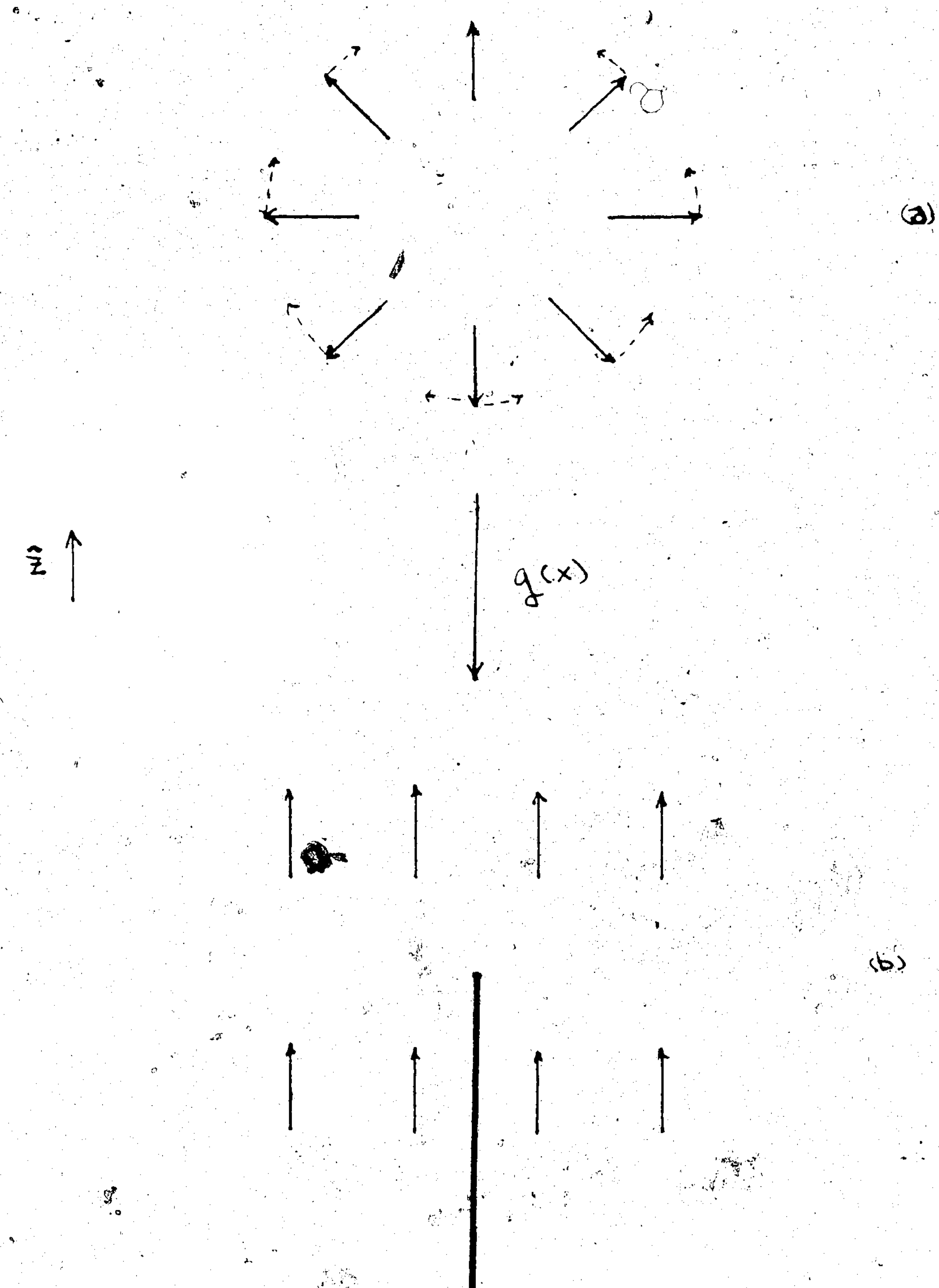


Figure 9. The gauge transformation of the monopole into the Abelian gauge. The dotted lines in (a) indicate the direction of the rotation of the Higgs vectors required to reach the configuration in (b). These rotations clearly become undefined along the negative z axis.

In this gauge, the vector field possesses a Dirac string singularity exactly as in conventional electrodynamics. This string appears to compensate for the loss of the topological configuration carried by the Higgs field and allows the magnetic charge, which is inherently topological, to be carried by the gauge field. Explicitly

$$g_M = \frac{1}{4\pi} \int_{S^2} d\sigma_i \, \varepsilon_{ijk} M^{jk} = \frac{1}{e}$$

The mass of the monopole may be identified with the energy obtained from the energy integral

$$E = - \int L(x) d^3x = \int d^3x [\tfrac{1}{4} G^a_{ij} G^{aij} - \tfrac{1}{2} D_i \phi^a D^i \phi^a + \frac{\lambda}{4} (\phi^a \phi^a - F^2)^2]$$

Inserting the asymptotic fields (1.27) yields an infinite result, since these forms of the fields do not incorporate the correct boundary properties as $r \to 0$. Modifications to (1.27) occur inside the Compton wavelength of the heavy gauge bosons. Approximate numerical values for E have been obtained by 't Hooft and others, however, and these give

$$M \approx 137 \, M_W$$

where M_W is the mass of the heavy vector particles. This shows that for realistic compact gauge models, the mass of the monopole would be enormous and its experimental absence is not surprising.

CHAPTER II

TOPOLOGICAL CONFIGURATIONS IN EUCLIDEAN SPACE

2.1 Introduction

Having considered in Chapter I the possibility of topological solitons in the classical version of unified gauge models of the weak and electromagnetic interactions, the next step would logically be to look for analogous solutions in quantum chromodynamics[7]. This is especially true since quantum chromodynamics is expected to contain certain important non-perturbative aspects. In exploring this possibility, however, the topological analysis of Chapter I seems to immediately imply negative results. First of all, quantum chromodynamics is an unbroken theory and one does not want to introduce the Higgs scalars which have been present in the previous models which possessed soliton solutions.

This problem may be circumvented, however, since the classical vacuum, defined by $F^a_{\mu\nu} = 0$, in general includes fields which are non-vanishing. These are just the gauge transforms of $A^a_\mu = 0$, given by

$$A^a_\mu = \frac{i}{g} (g^{-1}(x)\partial_\mu g(x))^a \qquad g\epsilon G$$

Thus if the mappings, $g(x)$, from the boundary of space into the group manifold fall into non-trivial homotopy classes, pure gauge field solitons may be expected. In other words, for three dimensional space, if $\Pi_2(G)\neq\{e\}$, the space of finite energy gauge field configurations will fall into distinct sectors determined by the topology of the pure gauge field at the boundary. Lumps of gauge field energy ("glueballs" for Q.C.D.) in a sector other than that with the trivial boundary condition $A^a_\mu=0$, could not dissipate. This follows due to an important connection between the topology of the gauge fields and the existence of non-vanishing $F_{\mu\nu}$ in the manifold (see below; e.g. equation 2.4).

Unfortunately, Table 1 of Chapter I shows that $\Pi_2(G)$ is trivial for every group of interest. Also, including fermion fields will not help since their vacuum expectation values must vanish and they therefore cannot contribute to the formation of topologically distinct vacuum sectors (ignoring the possibility of dynamical symmetry breaking via fermion bound states). This negative result from homotopy theory has also been verified by use of a scaling argument by Coleman[30] (note also (31)).

An examination of Table 1 does show, however, that $\Pi_3(G)$ is non-trivial for every G but U(1). Since $\Pi_3(G)$ represents the classes of maps from hyperspheres (S^3) into the group manifold and since four dimensional space is bounded by a hypersphere (S^3), the nontriviality of $\Pi_3(G)$ implies topologically stable solitons should exist in four space dimensions (energy conservation rules out a localized soliton in four dimensional Minkowski spacetime). Therefore a pure SU(3) or SU(N) Yang-Mills theory in Euclidean spacetime should possess soliton solutions labelled by the integers $Z=\Pi_3(SU(N))$.

At first, the existence of solitons in four dimensional Euclidean spacetime seems totally irrelevant to considerations in Minkowski spacetime. Polyakov[33][34], however, was the first to point out that the path integral formulation of field theory is (properly) defined in terms of integrals over classical field configurations in Euclidean space, each weighted by e^{-S} where S is the action, and the resulting Greens functions are then analytically continued into Minkowski space. Therefore classical Euclidean fields with stationary action could produce important contributions, non-perturbative in nature. In fact, he argued that in certain regimes (for instance the infrared region), these contributions could dominate the path integral and possibly provide an explanation for quark confinement. Although these speculations have yet to be confirmed in their entirety, they provide strong motivation for searching for these solutions and exploring their effects in a Yang-Mills theory.

At this point, however, it should be checked that the arbitrariness in choice of a pure gauge boundary condition, required by these solitons, actually does exist in the quantum theory. In the Schrodinger representation of quantum field theory, a physical state is represented by a wave functional $\Psi[A_\mu \ldots]$. Now, any gauge transform of A_μ (A_μ^g) serves to define an equally valid functional $\Psi[A_\mu^g \ldots]$. In particular, the vacuum functional $\Psi_{vac}[A_\mu = 0]$ may be equally well described by the functional $\Psi_{vac}[A_\mu^g]$, where A_μ^g is a gauge transform of $A_\mu = 0$. The requirement that physical states be gauge invariant (in particular the vacuum), may be met by forming physical states which are superpositions of gauge transforms

$$\Psi_{phys}[A_\mu] = \int [dg]\ \psi[A_\mu^g]$$

Clearly gauge fields with non-trivial topological boundary conditions must be included in this superposition. In particular, a time slice at the boundary for a Euclidean soliton (note equation 2.7), will give a gauge field which represents one (topologically non-trivial) contribution to the physical vacuum at that time†

$$\Psi_{vac} = \int [dg]\ \psi_{\substack{pure \\ gauge}}\ [A_\mu^g]$$

Evaluation of matrix elements of physical states such as described involves infinities in the functional integration due to the infinite contribution of the group integration [dg]. These factors may be subtracted out, however, in a consistent manner to yield finite results (see section 3.3). This procedure is equivalent to fixing a gauge.

At this point it might be thought that a more correct approach would be to specify physical states in terms of purely "physical" degrees of freedom, such as $A_\perp$ ($\perp$=transverse) by working in a gauge such as the Coulomb gauge. In this way, configurations which go to topologically non-trivial boundary conditions would apparently be avoided. For example, the

---

†See section 4.1 for more details on this point.

vacuum would be fixed as $\Psi_{vac}[A_\perp = 0]$ and the vacuum configuration required at $t = +\infty$ by a Euclidean soliton should be excluded. In fact, this is not the case and as pointed out by Gribov, even the Coulomb gauge possesses ambiguities. These ambiguities are in fact just the necessary degrees of freedom required to incorporate solutions with non-trivial topology[32]. The role of these solutions is much more transparent, however, in more general gauges.

2.2 The Instanton or Pseudoparticle

The preceding arguments lead to a consideration of pure, sourceless Yang-Mills theory in Euclidean spacetime. Since the homotopy arguments are identical for gauge group G = SU(2) or SU(3) it is simpler to work with SU(2) even though SU(3) is the group of interest for Q.C.D. The action for a pure Yang-Mills field is[12]:

$$S = -\tfrac{1}{4}\int F^a_{\mu\nu}\, F^{\mu\nu a}\, d^4x \quad , \text{metric}=(1,1,1,1) \qquad (2.1)$$

where
$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^{\mu a} + g f^{abc} A^b_\mu A^c_\nu$$

When the group index a is omitted, the fields are to be interpreted as contracted with group generators, i.e.

$$A_\mu = \frac{g}{i} A^a_\mu T_a ; \; F_{\mu\nu} = \frac{g}{i} F^a_{\mu\nu} T_a$$

$$\Rightarrow F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

The factors of g/i are to conform with mathematical convention and since they make the fields anti-hermitian, they will only be used for this chapter which is primarily formal. Under a local gauge transformation g(x),

$$A_\mu \to g^{-1}(x)A_\mu g(x) + g^{-1}(x)\partial_\mu g(x)$$

and $$F_{\mu\nu} \to g^{-1}(x)F_{\mu\nu}g(x)$$

For gauge group SU(2), $T_a = \sigma_a/2$ with σ_a the Pauli spin matrices and $[T_a,T_b] = i\varepsilon_{abc}T_c$, i.e. $f_{abc} = \varepsilon_{abc}$

Also $$\mathrm{Tr}(T^aT^b) = \tfrac{1}{2}\delta^{ab}$$

$$\mathrm{Tr}(T^aT^bT^c) = \tfrac{i}{4}\,\varepsilon^{abc}$$

With group indices contracted, the Yang-Mills field equations may be written in the compact form:

$$D_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu} + [A_\mu,F^{\mu\nu}] = 0 \qquad (2.2)$$

Defining the tensor dual to $F^{\mu\nu}$:

$${}^*F_{\mu\nu} = \tfrac{1}{2}\,\varepsilon_{\mu\nu\lambda\rho}\,F_{\lambda\rho}$$

implies the Bianchi identity:

$$D_\mu {}^*F^{\mu\nu} = 0 \qquad (2.3)$$

which follows from the definition of ${}^*F_{\mu\nu}$ in terms of A_μ.

Now, our soliton configurations are characterized by integer topological quantum numbers, corresponding to

$Z = \Pi_3(SU(2))$. These integers can be written in the form:

$$q = -\frac{1}{16\pi^2}\int \mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu})d^4x \qquad (2.4)$$

q is referred to as the Pontryagin index or winding number of the manifold (Here manifold means the direct product of the group manifold and the 4-D spacetime manifold, i.e. the principal fiber bundle over E_4[35]). The interpretation of this integral form of these integers will be discussed in somewhat more detail below.

The integrand of equation (2.4) can be written as a divergence:

$$\mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu}) = \partial_\mu K^\mu$$

with $$K^\mu = 2\,\varepsilon^{\mu\nu\gamma\delta}\mathrm{Tr}(A_\nu\partial_\gamma A_\delta + \frac{2}{3}A_\nu A_\gamma A_\delta) \qquad (2.5)$$

$$\Rightarrow q = -\frac{1}{16\pi^2}\int \partial_\mu K^\mu d^4x$$

or $$q = -\frac{1}{16\pi^2}\int_{S^3_\infty} K^\mu d^3\sigma_\mu \qquad (2.6)$$

where $d^3\sigma_\mu$ is a volume element on the hypersphere at infinity S^3_∞. Requiring the action to be finite imposes the restriction that $F_{\mu\nu}$ go to zero on the hypersphere at ∞, and therefore A_μ must be pure gauge:

$$A_\mu = g^{-1}(x)\partial_\mu g(x) \qquad (2.7)$$

where g(x) is an element of SU(2), or more correctly, g(x) associates an element of SU(2), for each value of x. Now equation (2.5) can be rewritten

$$K^{\mu} = \varepsilon^{\mu\nu\gamma\delta} \mathrm{Tr}\{A_{\nu}F_{\gamma\delta} - 2/3\, A_{\nu}A_{\gamma}A_{\delta}) \qquad (2.8)$$

Since $F_{\gamma\delta}$ must vanish on the hypersphere at infinity, using (2.7) gives (2.6) in the form:

$$q = \frac{1}{24\pi^2} \int_{S^3_\infty} \varepsilon^{\mu\nu\gamma\delta} \mathrm{Tr}\{(g^{-1}\partial_{\nu}g)(g^{-1}\partial_{\gamma}g)(g^{-1}\partial_{\delta}g)\} d^3\sigma_{\mu} \qquad (2.9)$$

This shows that the g(x), the mappings from $x \varepsilon S^3_\infty$ onto the SU(2) group manifold, determine the winding number or homotopy class of the gauge field, as expected. In fact the integrand of equation (2.9) is just the Jacobian of the change from the coordinates on S^3_∞ to $S^3_{SU(2)}$, i.e.:

$$q = \frac{1}{24\pi^2} \int_{SU(2)} d\mu(g)$$

where $d\mu(g)$ is the invariant measure on the group manifold[37]. The integral gives a factor $24\pi^2$ for an integration once over the group manifold, hence q measures the number of times the SU(2) group manifold is covered when x is integrated over S^3_∞ .

The integral expression (2.4) is a specific example of a general mathematical relation for what is usually referred to as the Pontryagin number or index of the manifold. Since, for Yang-Mills theories, the fiber bundle is normally complex, the correct topological invariant is in general the Chern number (equivalent to the Pontryagin index for SU(2)) . The Chern number for a given manifold may be found by integrating the corresponding Chern form C(M); M represents the manifold[35][36].

The expression for C(M) is:

$$C(M) = \det(1 + \frac{1}{2\pi}\,\Omega) \qquad (2.10)$$

where Ω is the curvature 2-form of the manifold. In Yang-Mills theories the components of the 2-form make up the field tensor, i.e.

$$\Omega_{YM} = F = \tfrac{1}{2} F_{\mu\nu} dx^{\mu} \wedge dx^{\nu}$$

where $\wedge$ represents the wedge product satisfying $dx^{\mu} \wedge dx^{\nu} = -dx^{\nu} \wedge dx^{\mu}$. C(M) can be expanded:

$$C(M) = C_0 + C_1 + C_2 \ldots$$

where C_0 is a 0-form, c_1 a 2-form, C_2 a 4-form, etc., and C_0 is always equal to 1. The series is finite since the base space of the bundle (of dimension n) cannot admit higher than an n-form.

Recalling that F is in a form contracted with the SU(2) generators, we can write it in the form

$$F = F^a \frac{T_a g}{i} = F^a \sigma_a \frac{g}{2i}$$

$$= \frac{g}{2i} \begin{bmatrix} F^3 & F - iF^2 \\ F^1 + iF^2 & -F^3 \end{bmatrix}$$

and

$$C(M) = \det \begin{bmatrix} 1+g/4\pi F^3 & g/4\pi(F^1 - iF^2) \\ g/4\pi(F^1+iF^2) & 1 - g/4\pi F^3 \end{bmatrix}$$

$$= 1 - \frac{1}{16\pi^2} (F^a \wedge F^a)$$

where multiplication of the F's must be interpreted as wedge products since the F's are forms. This shows that the second Chern form is non-zero. Integrating C_2 gives an expression for the Chern number q_2.

$$q_2 = - \frac{2}{16\pi^2} \int Tr(F \wedge F)$$

$$= - \frac{1}{16\pi^2} \int \tfrac{1}{2} Tr\ (F_{\mu\nu} F_{\gamma\delta}) \varepsilon^{\mu\nu\gamma\delta} d^4x$$

$$= - \frac{1}{16\pi^2} \int Tr(F_{\mu\nu} {}^{*}F^{\mu\nu}) d^4x$$

which corresponds with equation (2.4). This also shows (2.4) remains valid for any gauge group G. One can also check that in the case of the Nielsen-Olesen vortex, one gets from equation (2.10):

$$C(M) = \det(1 + \frac{iF}{2\pi})$$

$$= 1 + \frac{iF}{2\pi}$$

F is a closed form (i.e. the exterior derivative dF = 0) which implies it can be written as the exterior derivative of a 1-form A. Integration of $C_1 = iF/2\pi$ therefore gives:

$$q_1 = \frac{i}{2\pi} \int_{R^2} F = \frac{i}{2\pi} \int_{R^2} dA$$

$$= \frac{i}{2\pi} \oint_{\partial R^2} A = \frac{e}{2\pi} \oint_{S^1} A_\mu dx^\mu = \Phi \times \frac{e}{2\pi}$$

$$= n \quad \text{the winding number}$$

where to conform with Yang-Mills convention we have set $A = A_\mu dx^\mu \times e/i$. This shows that the topological classification of the vortex in Chapter I, was simply based on the 1st Chern number of the U(1) bundle over R^2.

Returning to the Yang-Mills case, it can also easily be shown that non-trivial g(x) on the boundary S^3_∞, implies a non-zero Yang-Mills field somewhere within the boundary of the Euclidean 4-space. This follows, since one may insert the gauge function g(x) into (2.9). If this gives a non-zero q, then equation (2.4) implies $F_{\mu\nu}$ must be non-zero somewhere in E^4. Thus, non-trivial boundary conditions imply the 4-dimensional analogy of a soliton as expected from our homotopy arguments mentioned earlier.

An example of a non-trivial g(x) is[38]:

$$g(x) = \frac{(x_4 - i\vec{x}\cdot\vec{\sigma})}{\sqrt{x_\mu x^\mu}} \quad \text{where } x_4 \text{ is Euclidean time and } \mu = 1,\ldots,4 \tag{2.11}$$

$$\text{and} \quad g^{-1}(x) = \frac{(x_4 + i\vec{x}\cdot\vec{\sigma})}{\sqrt{x_\mu x^\mu}}$$

$$\text{This gives} \quad A_\mu = g^{-1}(x)\partial_\mu g(x) = -\frac{2i\sigma^{\mu\nu}x_\nu}{x_\mu x^\mu} \tag{2.12}$$

where $\sigma^{\mu\nu} = \frac{1}{4i}[\sigma^\mu, \sigma^\nu]$ with $\sigma^\mu = (-i\vec{\sigma}, 1)$. Substitution of the g(x) of (2.11) into equation (2.9) shows that this choice of g(x) gives q = 1.

A solution of the Yang-Mills field equations with $F_{\mu\nu}$ regular over all space and with nontrivial boundary conditions, was first found by Belavin, Polyakov, Schwartz and Tyupkin[39]. They exploited the fact that the Bianchi identity (2.3) is identically satisfied merely from the definition of ${}^{*}F_{\mu\nu}$. Hence the self-duality condition

$$F_{\mu\nu} = {}^{*}F_{\mu\nu} \qquad (2.13)$$

is equivalent to the field equation (2.2). They looked for a solution with $q = 1$ and the boundary condition (2.12). This led them to consider the ansatz:

$$A^{\mu} = -2if(x^2)\,\frac{\sigma^{\mu\nu}x_{\nu}}{x^2}$$

where $f(x^2) \to 1$ as $|x|\to\infty$, and $f(0) = 0$ to cancel the divergence of $\frac{1}{x^2}$ required by boundary conditions of (2.12). Substitution of this ansatz into equation (2.13) leads to the equation for $f(x^2)$:

$$x^2\,\frac{df}{dx^2} - f + f^2 = 0$$

This equation has [illegible] solution:

$$f(x^2) = \frac{x^2}{x^2+\rho^2}$$

which gives

$$A^{\mu} = -\,2i\,\frac{\sigma^{\mu\nu}x_{\nu}}{(x^2+\rho^2)} \qquad (2.14a)$$

and

$$F^{\mu\nu} = \frac{4i\rho^2}{(x^2+\rho^2)^2}\,\sigma^{\mu\nu} \qquad (2.14b)$$

The parameter ρ determines the range over which A^{μ} differs significantly from the vacuum value (2.12), hence ρ is commonly referred to as the "size" of the instanton†.

---

†Belavin et al. dubbed the solution the pseudoparticle but 't Hooft introduced the name instanton because of the localized nature of the solution in time. Both terms are commonly used.

The solution (2.14) is also commonly written in a form introduced by 't Hooft:

$$A^a_\mu = \frac{2\,\eta_{a\mu\nu}x}{g(x^2+\rho^2)} \qquad (2.15a)$$

and

$$F^a_{\mu\nu} = \frac{4\,\eta_{a\mu\nu}\rho^2}{g\,(x^2+\rho^2)^2} \qquad (2.15b)$$

with 't Hooft's η tensor defined by[16]

$$\eta_{a\mu\nu} = \varepsilon_{oa\mu\nu} + \tfrac{1}{2}\,\varepsilon_{abc}\,\varepsilon_{bc\mu\nu}$$

An anti-instanton solution, with $q=-1$, may also be found by requiring $F_{\mu\nu}$ be anti-self-dual and requiring $f(0)=1$ and $f(\infty)=0$.

The action associated with the instanton can be simply found by exploiting the inequality:

$$-\tfrac{1}{4}\int d^4x\,Tr(F^{\mu\nu}-{}^*F^{\mu\nu})^2 \geq 0 \qquad (2.16a)$$

or

$$-\tfrac{1}{4}\int d^4x\,Tr(F^{\mu\nu}F_{\mu\nu}) + \tfrac{1}{2}\int d^4x\,Tr(F^{\mu\nu}F_{\mu\nu}) - \tfrac{1}{4}\int d^4x\,Tr({}^*F_{\mu\nu}{}^*F_{\mu\nu}) \geq 0 \qquad (2.16b)$$

Noting that the action S is

$$S = -\frac{1}{2g^2}\int d^4x\;Tr(F^{\mu\nu}F_{\mu\nu})$$

and

$$F^{\mu\nu}F_{\mu\nu} = {}^*F_{\mu\nu}{}^*F^{\mu\nu} ,$$

equation (2.16b) becomes

$$S \geq \frac{8\pi^2}{g^2}\,q \qquad (2.17)$$

Now (2.16a) shows that when $F_{\mu\nu}$ is self-dual, the inequality becomes an equality and the action for the instanton is $8\pi^2$. It can also be seen that the instanton represents a minimum action configuration for fields with non-trivial topology. (2.17) also shows the feature characteristic of solitons that the action is inversely proportional to the coupling.

For the Yang-Mills field in Euclidean space, the stress-energy tensor can be written(34):

$$\theta_{\mu\nu} = -\frac{1}{4g^2}\mathrm{Tr}\{(F_{\mu\lambda}-{}^*F_{\mu\lambda})(F^{\nu\lambda}+{}^*F^{\nu\lambda})+(F_{\nu\lambda}-{}^*F_{\nu\lambda})(F^{\mu\lambda}+{}^*F^{\mu\lambda})$$

This form of $\theta_{\mu\nu}$ clearly shows that self-dual (or anti-self-dual) field configurations will have vanishing stress energy. This property of the instanton supplies us with a hint that the instanton will be important in considering the vacuum in Yang-Mills theories (Section 4.1).

2.3 Multi-Instanton Configurations

Atiyah and Ward(41) have shown that the instanton is the only solution to the self-duality constraints for q=1. This leads one to the consideration of field configurations with q>1. These solutions can be interpreted as multi-instanton configurations. The key to finding such solutions is in making a judicious choice of Ansatz for A . A convenient choice turns out to be (42)(38):

$$A_\mu = i\,\bar{\sigma}^{\mu\nu} a_\nu \qquad (2.18)$$

with $\bar{\sigma}^{\mu\nu} = -{}^*\bar{\sigma}^{\mu\nu}$

and $\bar{\sigma}^{ij} = \sigma^{ij}$ and $\bar{\sigma}^{i4} = -\bar{\sigma}^{4i} = -\sigma^{i4}$

The self-duality condition:

$$F_{\mu\nu} = {}^*F_{\mu\nu}$$

then becomes:

$$\partial_\mu a_\nu - \partial_\nu a_\mu = -\tfrac{1}{2}\,\epsilon_{\mu\nu\gamma\delta}(\partial_\gamma a_\delta - \partial_\delta a_\gamma) \qquad (2.19a)$$

and

$$\partial_\mu a^\mu + a_\mu a^\mu = 0 \qquad (2.19b)$$

(2.19a) is clearly satisfied if a is written as a gradient

$$a_\mu = \partial_\mu(\ln\chi) \qquad (2.20)$$

and therefore

$$A_\mu = i\bar{\sigma}^{\mu\nu}\partial_\nu(\ln\chi) \qquad (2.21)$$

Equation (2.19b) now takes the simple form:

$$\partial_\mu\partial^\mu\chi = \Box\chi = 0 \qquad (2.22)$$

The Pontryagin index may also be expressed in a simple form:

$$q = -\frac{1}{16\pi^2}\int d^4x\,\Box\,\Box\,(\ln\chi) \qquad (2.23)$$

A solution for equation (2.22) was found by 't Hooft[44]:

$$\chi = 1 + \sum_{i=1}^{n}\frac{\rho_i^2}{(x-y_i)^2} \qquad (2.24)$$

This has the simple physical interpretation of representing n instantons with sizes ρ_i and positions y_i, hence the n instanton solution is characterized by 5n parameters. Also, it can be verified that (2.23) gives q=n for χ given by (2.24). Jackiw, Nohl and Rebbi have pointed out however that this is not the most general form of χ. A solution to (2.22) depending on 5n+4 parameters may be written[43][38]:

$$\chi = \sum_{i=1}^{n+1} \frac{\rho_i^2}{(x-y_i)^2} \tag{2.25}$$

The additional 4 parameters do not have such a simple physical interpretation and in fact they are not physically relevant in all cases. In particular, for n=1, the only relevant parameters are the five corresponding to the position and size of the instanton.

The solution (2.25) has the feature that it results in A_μ and $F_{\mu\nu}$ which are invariant under the full conformal group (minus inversions), if conformal transformations are coupled with suitable gauge transformations. That is conformal transformations generate new A_μ and $F_{\mu\nu}$ corresponding to the same Pontryagin index and gauge equivalent to the old A_μ and $F_{\mu\nu}$.

These conformally invariant, multi-instanton configurations, are not in fact the most general. It has been shown that the most general form of the multi-instanton solution depends on 8n-3 parameters[46]-[48]. These have a simple physical interpretation as 4 space-time, 1 scale and 3 group orientation parameters for each instanton minus 3 parameters corresponding to an irrelevant global group orientation. Belavin and Zacharov have obtained a form of this more general solution utilizing the inverse scattering method[44]. A summary of more mathematical approaches to this problem may be found in reference (50).

2.4 The Meron

Although the instanton represents the only solution to the Yang-Mills field equations with Pontryagin index equal to one, it is interesting to consider a certain class of configurations with $q=\frac{1}{2}$ which satisfy the field equations "almost everywhere". These are the so-called merons†. More correctly, we consider configurations with q=0 which correspond to pairs of merons and anti-merons.

The general form of the meron solution is[51]:

$$A_\mu(x) = \tfrac{1}{2}\, g^{-1}(x)\partial_\mu g(x) \tag{2.26}$$

and

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \\ &= -[A_\mu, A_\nu] \end{aligned} \tag{2.27}$$

This last line indicates that the non-vanishing of $F_{\mu\nu}$ is due to A_μ being $\frac{1}{2}$ of a pure gauge field. This also shows that the meron has $F_{\mu\nu}\neq 0$, over all space and hence infinite action. The Pontryagin index is given by:

$$\begin{aligned} q &= -\frac{1}{16\pi^2}\int d^4x\,\mathrm{Tr}(F_{\mu\nu}{}^*F^{\mu\nu}) \\ &= -\frac{1}{32\pi^2}\int d^4x\,\mathrm{Tr}\{([A_\mu, A_\nu][A_\gamma, A_\delta]\varepsilon^{\mu\nu\gamma\delta}) \end{aligned} \tag{2.28}$$

† Although these configurations were first discussed by de Alfaro, Fubini and Furlan, the name meron was coined by Callan, Dashen and Gross; the word meron comes from the Greek word for fraction.

Now, noting that $[A_\mu,A_\nu]^a = \epsilon_{abc}A^b_\mu A^c_\nu + \frac{1}{2}\, g^{-1}(x)[\partial_\mu,\partial_\nu]g(x)$, it can be seen that the term in brackets in (4.3) will vanish except where the gauge function is singular. Also, since the form (2.26) for A_μ will imply singularities one cannot utilize the form (2.9) for q. A choice of g(x) equivalent to that specifying the instanton boundary condition:

$$g(x) = \frac{x_4 - i\vec{\sigma}\cdot\vec{x}}{x_\mu x^\mu} \qquad (2.29)$$

leads in (2.26) to:

$$A_\mu = -i\,\frac{\sigma_{\mu\nu}x^\nu}{x_\mu x^\mu} \qquad (2.30)$$

The gauge function of (2.29) is singular in the sense that $g^{-1}[\partial_\mu,\partial_\nu]g \neq 0$, at 0 and infinity. Now, since the action density $S(x)$ and Pontryagin charge density $Q(x)$ (= integrand of (2.28)) are conformally invariant, it is convenient to utilize a conformal transformation to shift the singularities to two arbitrary points a and b. The gauge function g(x) now has the form:

$$g(x) = \frac{\bar{\sigma}^\mu(x-a)_\mu}{|x-a|}\cdot\frac{\sigma^\gamma(x-b)_\gamma}{|x-b|}\cdot\frac{\bar{\sigma}^\delta(a-b)_\delta}{|a-b|}$$

and the topological charge density has the form:

$$Q(x) = \frac{1}{2}\,(\delta(x-a) - \delta(x-b))$$

The topological singularities at a and b are termed merons and anti-merons respectively.

In a similar manner, it is possible to interpret configurations with topological charge equal one as systems of two merons and the instanton may be deformed so that its smooth topological charge density Q(x) dissociates into two points, each with $q=+\frac{1}{2}$.

At this point, one might ask whether integer topological charge may be dissociated in this manner into an arbitrary number of fractional charges. This, however, does not appear to be possible in any well defined manner for $q < \frac{1}{2}$. For a detailed discussion of this point, see reference (56).

2.5 Discussion

As mentioned in the introduction to this chapter, examining topological field configurations in Euclidean space might provide important non-perturbative information about the Minkowski space theory. It should be noted, therefore, that this "trick" applies equally well to the solitons of Chapter I if one looks at lower dimensional theories. For example, the vortex of section 1.3 may be considered an instanton in the + dimensional Abelian Higgs model and can be shown to have dramatic effects in that model[52][53]. Similarly, the monopole of section 1.5 may be considered as an instanton in compact quantum electrodynamics in 2+1 dimensions, and Polyakov has in fact shown that these "instanton" contributions lead to charge confine[illegible]hat model[34]. Such considerations can serve to give valuable insights into the rôle of the instanton in the much more intractable case of 3+1 dimensional quantum chromodynamics.

It should also be noted that instanton effects may be relevant for the Weinberg-Salam model of weak and electromagnetic interactions. For that model G/H=SU(2) and therefore $\Pi_3(G/H)=Z$ and topological solitons in Euclidean space may exist. Because of the Higgs scalars the equations are more complex than those of the pure Yang-Mills case and an interpretation also correspondingly more difficult[16].

CHAPTER III

QUANTUM THEORY OF SOLITONS

3.1 Introduction

Up to this point, all the considerations have been purely classical. At least classical in the sense that all the field equations considered have been treated as c-number equations. The relevance of the preceding analysis thus depends on the role which the topological solitons which have been found, play in a quantized theory. Although all the problems associated with this question have not yet been resolved, considerable insight has been gained through a variety of approaches (e.g. (57)-(64)). Reference (57) utilizes a generalization of the W.K.B., semi-classical approximation; reference (58) a self-consistent formulation of soliton matrix elements; reference (60) utilizes canonical quantization methods and reference (62) utilizes the boson method. Of these approaches the most consistent is probably that of Umezawa et al.[(60)], since it begins with a quantum field theory and generates soliton solutions in the tree approximation via the boson method. In all these approaches, however, the generalization of the methods to include instantons is extremely difficult.

Here the path integral formulation of field theory will be used to incorporate solitons and instantons. This approach has the advantage of giving an extremely intuitive insight into the role of solitons, and instantons in particular and also follows the current trend toward using the path integral method in quantizing gauge theories.

In this chapter the path integral formulation of quantum mechanics and quantum field theory will be outlined and the modifications required for the introduction of solitons described. The essential modification is the introduction of collective coordinates to describe the location of the soliton[63][64]. With a few alterations[72][75], these collective coordinates allow a consistent treatment of the instanton with the path integral now in Euclidean space.

3.2 Path Integral Quantization

The path integral formalism developed by Feynman provides an extremely elegant formulation of quantum mechanics which can be derived directly from the standard formulation of quantum mechanics[12][65] or motivated directly from first principles[66]. Here we will merely outline the formalism.

In the path integral approach to quantum mechanics, the transition amplitude between a single particle state $|q,t\rangle$ and another state $\langle q',t'|$, is given by the expression:

$$\langle q',t'|q,t\rangle = \langle q'|e^{-iTH/\hbar}|q\rangle = \int \left[\frac{dq}{\sqrt{2\pi\hbar}}\right]\left[\frac{dp}{\sqrt{2\pi\hbar}}\right]\exp\left[\frac{i}{\hbar}\int_t^{t'} (p\dot{q} - H(p,q))d\tau\right] \quad (3.1)$$

where $T = t'-t$.

That is, the total quantum mechanical amplitude is just a sum over all possible paths in phase space, with each path weighted by the action (S) for the path, times $i/\hbar$.

The integral over the functions q(t), p(t) must be understood as an infinite product of ordinary integrals at fixed times†. That is, one divides the time interval T into n+1 equal segments with spacing ε i.e. $t_i - t_{i-1} = \varepsilon$. Equation (3.1) is thus more correctly written:

$$\langle q',t'|q,t\rangle = \lim_{h\to\infty} \int \prod_{i=1}^{n} dq_i \prod_{i=1}^{n+1} \frac{dp_i}{2\pi\hbar} \exp\{\frac{i}{\hbar} \sum_{j=1}^{n+1} [p_j(q_j - q_{j-1}) - H(p_j, \tfrac{1}{2}(q_j - q_{j-1}))(t_j - t_{j-1})]\} \qquad (3.2)$$

thus $\quad \int [\frac{dq}{\sqrt{2\pi\hbar}}][\frac{dp}{\sqrt{2\pi\hbar}}] \equiv \int \prod_{\tau} \frac{dq(\tau)dp(\tau)}{2\pi\hbar}$

The factors of $\hbar$ have been explicitly exhibited in these expressions but from here on we will set $\hbar=1$ again, and absorb the factors $1/\sqrt{2\pi\hbar}$ into dq and dp. The path integrations in (3.1) and (3.2) must be performed with the boundary conditions $q \to q'$ as $t \to t'$ and $q \to q$ as $t \to t$.

Equation (3.1) immediately generalizes to the case with n particles (non-interacting) or 2n degrees of freedom described by q_n, p_n

$$\langle q_1', q_2' \ldots q_n', t'|q_1, q_2, \ldots q_n, t\rangle = \int \prod_{n=1}^{N} [dq_n dp_n] \exp \{i\int_t^{t'} \sum_{n=1}^{N} p_n q_n - H(p_i, q_i)]d\tau\} \qquad (3.3)$$

† For a careful mathematical treatment of integration over function spaces, see (67).

Now the only path integrals which one can do explicitly are generalizations of the simple Gaussian integral

$$\int_{-\infty}^{\infty} e^{-\alpha^2 x^2/2}\, dx = \sqrt{\frac{2\pi}{\alpha}}$$

The generalization of this to an N dimensional integral is

$$\iiint_{-\infty}^{\infty} \ldots \int dx_1 \ldots dx_N \exp\{-\tfrac{1}{2} \sum_{ij} x_i K_{ij} X_j\} = (2\pi)^{N/2} (\det K)^{-\frac{1}{2}} \qquad (3.4)$$

where K is some matrix. This can be easily checked (for K diagonal) to follow from the one-dimensional Gaussian integral. A somewhat more general form of (3.4), which will be useful later, is

$$\iiint dx_1 \ldots dx_N \exp\{-\tfrac{1}{2} \sum_{ij} x_i K_{ij} x_j + \sum_i a_i x_i\}$$
$$= (2\pi)^{N/2} (\det K)^{-\frac{1}{2}} \exp\{\tfrac{1}{2} \sum_{ij} a_i K^{-1}_{ij} a_j\} \qquad (3.5)$$

If the argument of the exponent in (3.2) can be written as a simple quadratic form in p or q, equation (3.4) can be applied immediately to the expression (3.2) for the path integral. Equivalently, one can consider a continuum generalization of (3.4) with K now in general a differential operator, and apply it (carefully) directly to the compact form of the path integral (3.1) or (3.3). The factors of (2Π) can be seen to cancel the factors of $(2\pi)^{-1}$ chosen in the integration measure.

A trivial example is the case where the Hamiltonian takes the form:

$$H = \frac{p^2}{2} + V(q)$$

The p functional integration in (3.2) can now be performed using (3.4) to give:

$$<q't'|qt> = \int [dq]\exp\{i \int_t^{t'} L(q,\dot{q})d\tau\} \quad (3.6)$$

with $\int L(q,\dot{q})d\tau$ = action = $S[q]$ and where

$$L(q,\dot{q}) = \tfrac{1}{2}\dot{q}^2 - V(q) = \text{Lagrangian}$$

Here equations (3.2) and (3.5) have been used in the form

$$\int[dp]\, e^{i\int\{-p^2/2+p\dot{q}\}d\tau} = e^{i\int\frac{1}{2}\dot{q}^2 d\tau}$$

The path integral is often expressed in the simpler form (3.6), and in fact this was the form originally introduced by Feynman. It is not, however, valid for a velocity dependent potential $V(q,p)$.

The expression (3.3) allows for a straightforward generalization to field theory $(q\to\phi,\ n\to x)$ (12). For a single scalar field ϕ, the time evolution operator evaluated between two states, $|i>$ and $<f|$ is:

$$<f|e^{-iTH}|i> = \int [d\phi][d\Pi]\exp\{i \int_t^{t'} d\tau \int d^3x[\Pi(x,\tau)\frac{\partial\phi}{\partial\tau}(x,\tau) - \mathcal{H}(x,\tau)]\} \quad (3.7)$$

with $\Pi = \dfrac{\partial\mathcal{L}}{\partial\dot{\phi}}$, $\tau = t' - t$ and $\mathcal{H}$ = Hamiltonian density.

This expression must also be understood as a compact notation for an infinite product of integrals at fixed time and now also fixed points in space. This is done by dividing space into cells of volume ε^3, labelled by an index α. Equation (3.7) is thus in essence a limit of equation (3.3) as $N\to\infty$ with q identified with ϕ and $n(\equiv\alpha)$ describing the x dependence of ϕ(12). Again, boundary conditions must be correctly incorporated into the path integral.

A convenient way to make contact with perturbation theory is to use the path integral to represent Schwinger's generating functional Z[J](54). Z[J] is basically the vacuum to vacuum transition amplitude in the presence of an external source J(x). In the path integral formalism Z[J] is given by

$$Z[J] = N \int [d\phi][d\Pi]\exp\{i \int_{-\infty}^{+\infty} d^4x[\Pi(x)\dot{\phi}(x)-H(x) + J(x)\phi(x)]\} \tag{3.8}$$

In this expression the limit $t'\to+\infty$, $t\to-\infty$ has been taken of the form given in equation (3.7). N is a normalization factor, which will never have to be explicitly evaluated, which basically ensures $Z[J] \equiv \langle 0|0\rangle_J \to 1$ as $J\to 0$. Now the Green's functions, and hence all the information contained in the theory, may be obtained by repeated functional differentiations of Z[J] with respect to J.

$$G(x_1 \ldots x_N) = (i)^{-n} \left. \frac{\delta^n Z[J]}{\delta J(x_1)\delta J(x_2)\ldots\delta J(x_n)} \right|_{J=0} \tag{3.9}$$

where $G(x_1 \ldots x_n)$ represents the complete n-point Green's function defined by

$$G(x_1 \ldots x_n) = \langle 0|T(\phi(x_1)\phi(x_2)\ldots\phi(x_n))|0\rangle$$

and where T denotes time ordering. These Green's functions contain contributions from all Feynman diagrams including disconnected vacuum to vacuum graphs. The connected Green's functions are given by

$$G_c(x_1 \ldots x_n) = (i)^{-n} \frac{1}{Z[J]} \frac{\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0} \quad (3.10a)$$

Equivalently one can define:

$$Z[J] = \exp\{iW[J]\}$$

and then

$$G_c(x_1 \ldots x_n) = (i)^{1-n} \frac{\delta^n W[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0} \quad (3.10b)$$

S matrix elements may be obtained from these connected Green's functions by applying the L.S.Z. reduction formula, i.e.

$$S_{fi} = \langle p_1 \ldots p_n; out | q_1 \ldots q_n; in \rangle$$

$$= \left(\frac{i}{\sqrt{Z}}\right)^{m+n} \prod_{i=1}^{m} \int d^4x_i \prod_{j=1}^{n} \int d^4y_j \; f_{q_i}(x_i)(\overrightarrow{\Box_{x_i} + m^2})$$

$$G_c(y_1 \ldots y_n, x_1 \ldots x_m) \times (\overleftarrow{\Box_{y_j} + m^2}) f^*_{p_j}(y_j) \quad (3.11)$$

The meaning of $\sqrt{Z}$ (not to be confused with Z[J]) and f_{q_i}, are explained in reference (68), Chapter 16.

To give a well defined meaning to the path integral in equation (3.8), the integration is normally performed in **Euclidean space (x_4=-ix_o) and the resulting Green's functions are analytically continued back into Minkowski space by a**

Wick rotation. It has been proven that if the fields in Euclidean space satisfy certain axioms, this procedure is well defined (12). It can also be seen that in general it is not necessary to specify boundary conditions on ϕ in the functional integration in equation (3.8). This is true because (e.g. in Euclidean space), as the time goes to infinity, all configurations suffer an infinite damping factor except that with zero energy; the vacuum. That is

$$Z[J]_T = \langle 0|e^{-HT}|0\rangle_J = \langle 0|e^{-\infty}|0\rangle_J \quad \text{as} \quad T \to \infty$$

unless $H|\phi\rangle = 0$. This argument only applies, however, when the fields ϕ are not separated from the vacuum by a superselection rule (or topological quantum number).

Most field theories have a form which allows equation (3.8) to be simplified. That is, if the Hamiltonian density can be written in the form

$$H(x) = \tfrac{1}{2}\,\Pi^2(x) + f[\phi(x),\vec{\nabla}\phi(x)]$$

then the Π integrations in (3.8) may be carried out trivially to give

$$Z[J] = N \int [d\phi]\exp\{i \int [L(x) + J(x)\phi(x)]d^4x\} \qquad (3.12)$$

with L the Lagrangian density given by:

$$L(x) = \tfrac{1}{2}\,(\dot{\phi})^2 - f[\phi(x),\vec{\nabla}\phi(x)]$$

Equation (3.12) is the expression normally given for the generating functional. The normalization factor N is used to absorb any constants from the expression for Z[J] since equation (3.10a) shows that these will be irrelevant when computing physical quantities.

Equation (3.12) (or (3.8)) can be used to develop a perturbation expansion for the Green's functions in a straightforward way. One can write the action as:

$$S[\phi] = \int L\, d^4x = \tfrac{1}{2} \int \phi(x)K(x,y)\phi(y)\, d^4x d^4y + \int L_I d^4x \quad (3.13)$$

The first term corresponds to the free field part of the action

$$S_o = \int L_o d^4x = \tfrac{1}{2} \int [(\partial_\mu\phi)^2 - \mu^2\phi^2]d^4x \quad \text{(for scalar field)}$$

and therefore

$$K(x,y) = (-\partial^2 - \mu^2)\delta^4(x-y) \quad (3.14a)$$

Again, equations (3.13) and (3.14) must be understood as limiting forms of expressions defined on a discrete space-time lattice. The propagator $\Delta(x,y) \equiv K^{-1}(x,y)$ may be defined by:

$$\int \{K(x,z)\Delta(z,y)\} d^4z = \delta^4(x-y)$$

or

$$(-\Box_x - \mu^2)\Delta(x,y) = \delta^4(x-y); \Box_x \equiv \partial_x^2 \quad (3.14b)$$

Also, the source term in equation (3.12) allows us to rewrite (3.12)

$$Z[J] = N\int[d\phi]\exp\{i\int L_I(\tfrac{1}{i}\ \tfrac{\delta}{\delta J(x)})d^4x\}\exp\{i\int\{L_o+J(x)\phi(x)]d^4x\}$$

$$= N\exp\{i\int L_I(\tfrac{1}{i}\ \tfrac{\delta}{\delta J})d^4x\}\int[d\phi]\exp\{i\int[L_o(\phi)+J(x)\phi(x)]d^4x\} \quad (3.15)$$

One can now do the path integration over ϕ explicitly by applying equation (3.5) to the discrete form of the ϕ integration in (3.15) and then taking the continuum limit. More transparently, the continuum generalization of equation (3.5)

$$\int[d\phi]\exp\{-\tfrac{1}{2}\int[\phi(x)K(x,y)\phi(y)d^4xd^4y]+\int J(x)\phi(x)d^4x\}$$

$$= (\det K)^{-\frac{1}{2}}\exp\{\tfrac{1}{2}\int J(x)K^{-1}(x,y)J(y)d^4xd^4y\} \qquad (3.16)$$

may be applied directly to (3.15) to give

$$Z[J]=N\exp\{i\int L_I(\frac{1}{i}\frac{\delta}{\delta J})d^4x\}\exp\{-\tfrac{1}{2}i\int J(x)\Delta(x,y)J(y)d^4xd^4y\} \qquad (3.17)$$

where the factor $(\det K)^{-\frac{1}{2}}$ has been absorbed into N. Expanding the first exponential:

$$\exp\{i\int L_I(\frac{1}{i}\frac{\delta}{\delta J(x)})d^4x\} = \sum_{n=0}^{\infty}\frac{(i)^n}{n!}[\int d^4x L_I(\frac{1}{i}\frac{\delta}{\delta J(x)})]^n$$

in equation (3.17), allows one to generate a Feynman graph perturbative expansion of Z[J]. Equation (3.10a) then allows one to compute the perturbative Green's functions to any order (ignoring renormalization problems).

The preceding analysis extends in a straightforward way to fermions if one utilizes results of functional integration over anti-commuting c-numbers[(70)(71)]. The relevant formula is,

$$\int [dx][dy]\exp\{\sum_{i,j} x_i B_{ij} y_j\} = \det B \qquad (3.18)$$

If one identifies x with the fermi field c-number ψ and y with its Hermitian conjugate $\psi^\dagger$ equation (3.17) can be used to incorporate fermions into the formalism described above. For example, in continuum notation equation (3.17) generalizes to

$$\int[d\psi^\dagger][d\psi]\exp\{\int \psi^\dagger(x)B(x,y)\psi(y)d^4xd^4y\} = \det B \qquad (3.19)$$

with B some diagonalizable differential operator.

Where the path integral can be done explicitly, the determinant in (3.19) or (3.16) is just the product of the eigenvalues of the corresponding differential operator. For example, for equation (3.16)

$$\mathrm{Det}\ K = \prod_n E_n \qquad (3.20a)$$

for E_n defined by

$$(\Box+\mu^2)\phi_n = E_n\phi_n \qquad (3.20b)$$

and ϕ may be expanded in terms of the ϕ_n .

3.3 Incorporation of Gauge Fields

The direct application of equation (3.12) of the last section with L a pure gauge field Lagrangian†, yields a meaningless infinity for the generating functional. This infinity results from functional integration over gauge equivalent paths (section 2.1). Therefore one must find a way to incorporate a gauge fixing term into the functional integral in a consistent way. This problem can be solved by

† The gauge field notations are now those of Chapter I, i.e. A_μ;$F_{\mu\nu}$ hermitian rather than anti-hermition, e.g. $A_\mu=A_\mu^a T^a$, etc..

utilizing an elegant method due to Fadeev and Popov[12][81][82].

The naive form of the generating functional is

$$Z[J] = N\int[dA_\mu]\exp\{i\int[L(A_\mu) + J_\mu A^\mu]d^4x\} \qquad (3.21a)$$

$$\equiv N\int[dA_\mu]\exp\{i(S[A_\mu]) + S_J)\} \qquad (3.21b)$$

If we represent an arbitrary infinitesimal gauge transformation by the operator

$$U(g) \simeq 1 + i\theta^a T^a + \ldots \qquad g\varepsilon G \qquad (3.22)$$

with $\theta^a(x)$ an arbitrary parameter and T^a the generators of G, an arbitrary gauge transform of A_μ is given by

$$A_\mu^g \equiv U(g)A_\mu U^{-1}(g) + \frac{i}{g}U(g)\partial_\mu U^{-1}(g) \qquad (3.23)$$

(the g in the denominator of the 2nd term is the coupling constant). One next defines a quantity Δ_{FP} by the expression

$$\Delta_{FP}[A_\mu] \int [dg(x)] \prod_{x,a} \delta[F^a(A_\mu^g(x))] = 1 \qquad (3.24)$$

Here $\prod_{x,a} \delta[\ ]$ represents an infinite product of delta functions defined at fixed x's (i.e. a delta functional) and

$$F^a(A_\mu(x)) = 0 \quad \text{(for all a)} \qquad (3.25)$$

represents an arbitrary linear gauge condition. F^a is in general an operator, e.g. for the Lorentz gauge $F^a(A_\mu) = 0$ represents $\partial_\mu A^\mu = 0$. Also, in equation (3.24), the integration over g(x) is interpreted:

$$\int [dg(x)] \Rightarrow \int \prod_a [d\theta^a(x)]$$

for θ^a defined by (3.22) and $g(x)=1$. Inserting the identity (3.24) into the R.H.S. of equation (3.1) yields

$$Z[J] = N\int[dA_\mu][dg(x)]\Delta_{FP}[A_\mu] \prod_{x,a} \delta[F^a(A_\mu^g)]\exp\{i(S[A_\mu]+S_J)\} \tag{3.26}$$

Since quantities of physical interest in (3.26) are gauge invariant (in particular $S[A_\mu]$, $\Delta_{FP}[A_\mu]$ and $[dA_\mu]$), a gauge transformation in the integrand of equation (3.26) will not affect Z[J]. A judicious choice of such a gauge transformation is

$$A_\mu(x) \to A_\mu^{g^{-1}}(x)$$

This removes the gauge dependence of A_μ^g and the entire integrand of (3.26) is independent of $g(x)$. The path integral $\int[dg(x)]$ can now be seen as an infinite factor independent of the fields. Absorbing this factor into N gives

$$Z[J] = N\int[dA_\mu]\Delta_{FP}[A_\mu] \prod_{x,a} \delta[F^a(A_\mu)]\exp\{i\int[L(x)+J^\mu(x)A_\mu(x)]d^4x\} \tag{3.27}$$

The path integral thus consists of a gauge fixing delta functional along with a Jacobian factor Δ_{FP} which corrects for changes of the gauge fixing condition $F^a(A_\mu) = 0$.

Some formal manipulations are required to put equation (3.27) in a tractable form. First of all, the delta functional in (3.27) may be rewritten[82]:

$$\prod_{x,a} \delta[F^a(A_\alpha)] = \exp\{-\frac{1}{2\alpha}\int[F^a(A_\mu)]^2 d^4x\} \tag{3.28a}$$

(α a real parameter)

Equal signs were not used since the two sides differ by a constant term which may be absorbed into N in equation (3.27). Next, it can be shown[82] that Δ_{FP} takes the form:

$$\Delta_{FP} = \det\left(\frac{\partial F^a}{\partial A_\mu^b} D_\mu^{bc} \delta^4(x-y)\right) \equiv \det M_{FP} \tag{3.28b}$$

where D_μ^{bc} is an element of the covariant derivative matrix,

$$D_\mu^{ab} = \delta_{ab}\partial_\mu + gf^{abc}A_\mu^c = (\partial_\mu - igA_\mu)^{ab}$$

The rule for functional integration over anti-commuting c-numbers given in equation (3.17) allows the determinant in (3.28b) to be written as a functional integral:

$$\Delta_{FP} = \int[dc^+][dc]\exp\{i\int c_a^+ \frac{\partial F^a}{\partial A_\mu^b} D_\mu^{bc} c_c\, d^4x\}$$
$$= \int[dc^+][dc]\exp\{i\int c_a^+[M_{FP}(x,y)]_{ab} c_b\, d^4x d^4y\} \tag{3.29}$$

The complex anti-commuting c-numbers, c^+ and c are the so-called Fadeev-Popov ghost fields. They are Lorentz scalars which obey Fermi-Dirac statistics, and are introduced merely as a convenient device for keeping track of the determinant Δ_{FP} in perturbation theory. The ghost fields serve to counterbalance the unphysical longitudinal modes of the vector field when one quantizes with a covariant (non-canonical) gauge fixing term. In quantum electrodynamics these fields are not required since the ghost fields do not couple to the photon field and the unphysical photon polarization states also decouple from the s-matrix.

Utilizing equations (3.28a) and (3.29), equation (3.27) for the generating functional takes the form:

$$Z[J] = \int[dA_\mu][dc^+][dc]\exp\{i[S_{eff} + \int J_\mu(x)A^\mu(x)d^4x]\} \quad (3.30)$$

where:

$$S_{eff} = \int[-\tfrac{1}{4}F^a_{\mu\nu}F^{\mu\nu a} - \frac{1}{2\alpha}(F^a[A_\mu(x)])^2 + c^+_a \frac{\partial F^a}{\partial A^b_\mu} D^{bc}_\mu c_c]d^4x$$

$$\equiv S_{\substack{gauge\\field}} + S_{\substack{gauge\\fix}} + S_{ghost}$$

Equation (3.10) may be used to generate a consistent perturbation theory in any gauge. Complete Feynman rules may be found in (12) and (7).

3.4 Incorporation of Solitons

The path integral approach to quantum field theory shows clearly that solitons will play a role in the full quantum theory. Since the path integral covers all classical field configurations, weighting each with a factor e^{-S} (in Euclidean space), and since solitons represent local minima of S, they can be expected to give an important contribution to the complete path integral. The exact nature of this contribution is not clear, however, since soliton field configurations do not satisfy the boundary conditions required of the generating functional and hence do not fit in a straightforward manner into the conventional perturbative approach.

A convenient manner in which to get some insight into the role of solitons in the quantum theory, is to use the stationary phase approximation. This gives a semi-classical approximation to the full quantum theory[4][57]. For the integral of a function f(x) with a local minimum at x=a; the stationary phase approximation gives the result

$$\int_{-\infty}^{\infty} dx\, g(x)e^{if(x)} = \int_{-\infty}^{\infty} dx\{g(a)+(x-a)g'(a)+\ldots\}\times\exp\{i[f(a) + \frac{(x-a)^2}{2} f''(a)+\ldots]\}$$

$$\simeq \sqrt{2\pi}\ \exp\{+if(a)\}g(a)\times(if''(a))^{-\frac{1}{2}}$$

That is, one expands f(x) and g(x) to second order in $\delta x \equiv x-a$. This approximation clearly depends on g(x) being a slowly varying function and on $f''(a)\neq 0$.

The generalization to an n-dimensional integral with $f(\vec{x})$ ($\vec{x}=(x_1,\ldots,x_n)$) having a local minimum at $\vec{x}=\vec{a}$, is:

$$\int dx_1,\ldots dx_n g(\vec{x})\exp\{if(\vec{x})\} \simeq I\exp\{if(\vec{a})\}g(\vec{a}) \qquad (3.31)$$

where $I = \int dy_1\ldots dy_n\ \exp\{\frac{i}{2}\sum_{ij}^{n} y_i A_{ij} y_j\}$

$$= (\frac{2\pi}{i})^{N/2}\ (\text{Det } A)^{-\frac{1}{2}}$$

where $A_{ij} = \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\vec{x}=\vec{a}}$ and $y_i = x_i - a_i$

If $f(\vec{x})$ has several local minima the integral gets a contribution from expanding about each minimum.

In field theory, the analogous procedure is to expand the action about local minima. That is, one writes the field near the minimum ϕ_o as

$$\phi(x) \simeq \phi_o(x) + \eta(x)$$

and then expands the action to second order in $\eta(x)$:

$$S[\phi] \simeq S[\phi_o] + \int \frac{\delta^2 S[\ ]}{\delta\phi(x)\delta\phi(y)}\Bigg|_{\phi=\phi_o} \eta(x)\eta(y)d^4x d^4y$$

The path integral may thus be approximated by:

$$<a|b> \doteq \int[d\phi]\exp\{iS[\phi]\}$$

$$\simeq \exp\{iS[\phi_o]\}\ (\text{Det } A)^{-\frac{1}{2}} \qquad (3.32)$$

where A is now given by:

$$A = \frac{\delta^2 S[\phi]}{\delta\phi(x)\delta\phi(y)}\Bigg|_{\phi=\phi_o}$$

If the Lagrangian takes the form

$$L = \tfrac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - U(\phi)$$

then the eigenvalue equation for A equivalent to (3.20b) is

$$\left(\frac{d^2}{dt^2} - \frac{d^2}{dx^2} + \frac{\partial^2 U}{\partial\phi^2}\Bigg|_{\phi=\phi_o}\right)\eta_n(x,t) = \lambda_n\eta_n(x,t) \qquad (x = \text{all space coordinates})$$

(3.33a)

where the perturbation $\eta(x,t)$ is expanded:

$$\eta(x,t) = \sum_n \eta_n(x,t) \qquad (3.33b)$$

In conventional field theory without solitons only fluctuations about the vacuum are considered. The action $S[\phi_o]$ is then zero and ϕ_o is normally set equal to zero. The application of the semi-classical, stationary phase approximation is then equivalent to a Feynman graph expansion, including graphs with one closed loop (e.g. reference (55)). To include solitons at this level would merely seem to require including contributions of the form of equation (3.32) with ϕ_o = soliton wavefunction. Two complications must be considered, however. The first is the requirement that field configurations included in the functional integral satisfy the proper boundary conditions. Since solitons possess non-trivial topology, they will in general only contribute to matrix elements between soliton states or soliton plus meson states. Such states may be referred to as being in the "soliton sector" whereas conventional states are in the "vacuum sector"[58].

The second problem concerns the existence of the translation mode discussed in Chapter I. The operator A has an n-fold degenerate (n=number of space dimensions) zero eigenvalue mode with eigenfunction:

$$\eta_{o_i} = \frac{\partial}{\partial x^i}(\phi_o) \qquad i=1,\ldots,n$$

where ϕ_o is the soliton solution corresponding to the translation mode discussed in section 1.2. This may be seen by writing

$$\eta_n(x,t) = \int d\nu\, e^{i\nu t}\psi_n(x)$$

Equation (3.33a) is then

$$[illegible]\frac{\partial^2}{dx[illegible]}\;[illegible]\frac{\partial[illegible]}{\partial\phi^2}\Big|_{\phi=\phi_o}[illegible]_n = \omega_n^2\,\psi_n$$

where

$$\omega_n^2 = \lambda_n + \nu^2$$

This equation is just the stability equation (1.8) which was shown to have an n-fold degenerate zero mode. Since ν^2 is positive, this also implies a zero eigenvalue λ_o for equation (3.33a). Since the determinant of A is given by:

$$\mathrm{Det}\ A = \prod_n \lambda_n$$

the existence of the zero eigenvalue λ_o will make Det A vanish. This will give an infinity in equation (3.32). This infinity is due to the arbitrariness of the location of the soliton:

$$\phi_o(x) \equiv \phi_o(x-x_o)$$

There is thus an arbitrariness about the choice of ϕ_o (i.e. choice of x_o) to perturb about and this gives rise to the infinity in the path integral. This problem does not arise when perturbing about the vacuum since the vacuum state is translationally invariant and unique.

The solution to this zero mode problem has been found within the semi-classical approximation to the path integral[57]. To allow perturbative calculations in the soliton sector to any order, however, it is necessary to consider a somewhat more general method; the method of collective coordinates[63][64]. In this approach, the position of the soliton x_o, is allowed to become a dynamical quantum variable $x_o(t)$ (i.e. one does a path integration over $x_o(t)$) referred to as a collective coordinate. Having introduced an extra degree of freedom into the path integral, it is necessary to introduce a constraint. If we write:

$$\phi(x,t) = \phi_o(x-x_o(t)) + \eta(\rho,t) \quad \text{where } \rho=x-x_o(t) \tag{3.34a}$$

a convenient constraint is:

$$\int \phi_o'(\rho)\eta(\rho,t)d^n\rho = 0 \qquad (\text{prime means } \frac{d}{d\rho}) \tag{3.34b}$$

This ensures that in the expansion of $\eta(x,t)$ in terms of the $\psi_n(x)$ as in equation (3.33b), the zero mode ψ_o is excluded. For the proper treatment of this transformation it is necessary to use the form (3.7) for the path integral. It is thus necessary to introduce a momentum $P(t)$ canonically conjugate to $x_o(t)$ and the momentum conjugate to $\eta(\rho,t)$ which we will denote $\tilde{\Pi}(\rho,t)$ and

$$\Pi(x,t) = \Pi_o(x-x_o(t)) + \tilde{\Pi}(\rho,t) \tag{3.35a}$$

The corresponding constraint is:

$$\int\phi_o'(\rho)\tilde{\Pi}(\rho,t)d\rho = 0 \tag{3.35b}$$

These constraints may be incorporated into the path integral by inserting delta functionals into the integrand.

The above canonical transformations may be justified more rigorously by inserting the identity[63]

$$\int[dx_o(t)][dP(t)]\delta[F_1(x_o(t),\phi)]\delta[F_2(P(t),\Pi,\phi)]\frac{\partial F_1}{\partial x_o}\frac{\partial F_2}{\partial P} = 1$$

directly into the original form of the path integral. The form of F_1 and F_2 are found by requiring η and $\tilde{\Pi}$, as defined by (3.34a) and (3.35a), to be canonically conjugate as well as $x_o(t)$ and $P(t)$. The resulting F_1 and F_2 are given by the L.H.S. of (3.34b) and (3.35b) respectively.

One must next reexpress the action $S[\phi,\Pi]$ in terms of the new functional integration variables $x_o(t)$, $P(t)$ (in general vectors in n-space) and $\eta(\rho)$ and $\tilde{\Pi}(\rho)$, taking into account the fact that the arguments of η and $\tilde{\Pi}$ have been shifted from x to $x-x_o\equiv\rho$. The result is[(63)(5)]

$$S[\phi,\Pi] = \int_t^{t'} \{\int \Pi\dot{\phi}\, d^n x - H[\phi,\Pi]\}dt = \int_t^{t'} L\, dt$$

$$= \int_{t_o}^{t'} \{P\dot{x}_o(t) + \int \tilde{\Pi}\dot{\eta}d\rho - H\}\, dt \qquad (3.36a)$$

where $$H = M_o + \overline{P}^2/2M_o + \int d\rho\{\tfrac{1}{2}\tilde{\Pi}^2 + \tfrac{1}{2}(\eta')^2 + U(\eta,\phi_o)\} \qquad (3.36b)$$

with M_o the soliton mass and

$$\overline{P} = \tfrac{1}{2}\{P(t) + \int d\rho\, \frac{\eta'(\rho,t)\tilde{\Pi}(\rho,t)}{(1+\xi/M_o)}\}$$

and $$U(\eta,\phi_o) = U(\phi_o+\eta) - U(\phi_o) - \left.\frac{dU}{d\phi}\right|_{\phi_o} \eta$$

$$\xi(t) = \int d\rho\, \phi_o'(\rho)\eta'(\rho,t)$$

Here primes denote differentiation wrt arguments (excluding t); dot means time differentiation and all ρ, x integrations are over all space dimensions.

The path integral is now:

$$\langle f|i\rangle = \int [dx_o][dP][d\eta][d\tilde{\Pi}]\delta[\int\phi_o'\eta d\rho]\delta[\int\phi_o'\tilde{\Pi}d\rho]\exp\{iS[x_o,P,\eta,\tilde{\Pi}]\} \qquad (3.37)$$

where $|i\rangle$ and $|f\rangle$ are in general states containing one soliton plus mesons.

To generate a perturbation series about soliton states one may introduce a generating functional for the one soliton sector in a way completely analogous to that in the vacuum sector.

$$Z_S[J,\tilde{J}] = N\int[dx_o][dP][d\eta][d\tilde{\Pi}]\delta[\int\phi_o'\eta d\rho]\delta[\int\phi_o'\tilde{\Pi}d\rho] \times$$

$$\exp\{i\int dt[\int(\tilde{\Pi}\dot{\eta}+J\eta+\tilde{J}\tilde{\Pi})d\rho - H]\}$$

$$\equiv \langle P'|P\rangle$$

where H is given by (3.36b) and P and P' are one soliton states. Now, if one writes H in the form

$$H = M_o + H_o + H_I$$

with
$$H_o = \tfrac{1}{2}\int d\rho[\tilde{\Pi}^2 + (\eta')^2 + \left.\frac{d^2U}{d\phi^2}\right|_{\phi_o}\eta^2]$$

then the path integral in η and $\tilde{\Pi}$ of H_o may be explicitly evaluated by expanding η and $\tilde{\Pi}$ in terms of the η_n of equation (3.33) (omitting η_o)(63). $Z_S[J,\tilde{J}]$ can therefore be written:

$$Z_S[J,J,x_o,P] = N\exp\{-i\int dtH_I(\frac{1}{i}\frac{\delta}{\delta J},\frac{1}{i}\frac{\delta}{\delta\tilde{J}})\}Z_o[J,J] \qquad (3.38)$$

where $$Z_o[J,\tilde{J}] = \int[d\eta][d\tilde{\Pi}]\delta[\int\phi_o'\eta d\rho]\delta[\int\phi_o'\tilde{\Pi}] \times$$

$$\exp\{i\int dt[\int d\rho(\tilde{\Pi}\dot{\eta} + J\eta + \tilde{J}\tilde{\Pi})d\rho - H_o]\}$$

where the exact form of $Z_o[J,\tilde{J}]$ depends on the form of $U(\phi)$. Equation (3.38) is completely analogous to equation (3.17) and can be used to generate a Feynman graph expansion in exactly the same manner. The expressions obtained, however, have a residual $x_o(t)$, $P(t)$ dependence which require a corresponding functional integration. Also, the perturbative expansion of Z_S brings out three kinds of propagators, that is, η-η, η-$\tilde{\Pi}$ and $\tilde{\Pi}$-$\tilde{\Pi}$ propagators corresponding to functional differentiations in H_I in (3.38) on $Z_o[J,\tilde{J}]$ wrt J-J, J-$\tilde{J}$ and $\tilde{J}$-$\tilde{J}$ respectively.

Using these methods, one can in principle calculate matrix elements between soliton plus meson states to any arbitrary accuracy. This allows one to calculate such things as the S-matrix for meson-soliton scattering to any order. Also, corrections to the soliton energy may be computed since

$$<P|e^{-i\int Hdt}|P> \equiv \exp[-i \int dt\ E(P)]$$

This may be evaluated perturbatively as in (3.38), with P treated approximately as a constant. Even for simple models, however, such calculations are quite complicated since the $\bar{P}^2$ term in equation (3.36) is non-polynomial in nature and as a result an infinite number of vertices appear in the Feynman graph expansion beyond the standard vertices found in the vacuum sector.

Explicit calculations have been done using the above methods for the case of the one-soliton sector of the ϕ^4 kink(63)(61). At the level of the graph corrections it was found that explicit Lorentz covariance, lost by introducing a soliton collective coordinate, was restored and

$$E(P) = (P^2 + M_o^2)^{\frac{1}{2}}$$

That is, the soliton behaves as a relativistic particle.

At the one loop level, renormalization is required and it turns out that the counterterms are the same as those in the vacuum sector. One loop corrections to the soliton mass may thus be found. It may also be checked that[63]

$$\langle P' | \phi(x,t) \, P \rangle$$

gives the soliton form factor with lowest approximation given by:

$$\langle P' | \phi(x) | P \rangle = \int dy \; e^{i(P-P')y} \; \phi_o(x-y) \quad (\text{for } t=0)$$

These results agree with similar calculations done in other quantization schemes[57][58]. At the two loop level and higher, however, an additional term enters H in equation (3.36) which is missed in the application of the canonical transformation as described above. This term is due to the necessity of keeping track of the order of the quantum operators while doing the canonical transformation to collective coordinates. This term becomes evident in the canonical quantization scheme[60] whereas it is obscured in the path integral approach where the variables are treated as c-numbers. The proper correction can be found within the path integral approach, however, by doing a more careful treatment of the transformation[64].

3.5 Incorporation of the Instanton

The instanton fits into the path integral formalism in a manner very similar to solitons in space dimensions. Collective coordinates are required for the zero modes resulting from the five arbitrary parameters which specify the instanton; that is, the four position coordinates and one scale parameter. In addition there exist zero modes corresponding to the possibility of arbitrary global gauge rotations. For SU(2) these require three additional collective coordinates. Integration over these three collective coordinates is trivial, however, and yields a simple numerical factor.

The four-dimensional nature of the instanton of course requires a modification of the procedure adopted for spatial solitons. Since the instanton degrees of freedom are covariant in nature, the collective coordinates may be introduced in an explicitly covariant manner and the path integral need not be formulated in terms of canonically conjugate variables. This amounts to a considerable reduction in complexity in the introduction of the collective coordinates. Also, since the path integration is in Euclidean space, the interpretation of the instanton is not as straightforward as for solitons in Minkowski space.

The quantity of interest is the vacuum† to vacuum amplitude in the presence of a single instanton, where

† As it turns out (section 4.1) the vacuum itself needs a careful definition for non-Abelian gauge theories.

initially the source term is omitted:

$$Z_1 \equiv \langle 0|0\rangle_1 = \frac{\int [dA_\mu][d\psi^\dagger][d\psi][dc^\dagger][dc]\exp\{-\int[L_{gauge}+L_{fermion}+L_{gauge\,fix}+L_{ghost}]d^4x\}}{\langle 0|0\rangle_0}$$

The normalization factor $\langle 0|0\rangle_0$ is the vacuum to vacuum amplitude with no instantons present.

Z_1 may be calculated to the one loop level using the stationary phase approximation(72)(73). That is, one expands each of the fields about its classical minima; the vector field about the instanton and the fermion and ghost field about $\psi=0$ and $c=0$. Writing

$$A = A_\mu^{(cl)} + A_\mu^{(qu)} \quad \text{where} \quad A_\mu^{(cl)} = \text{instanton}$$

the action becomes:

$$S[A,\psi C] \simeq S^{(cl)} + \tfrac{1}{2}A^{(qu)}M_A A^{(qu)} + \bar{\psi}M_\psi\psi + \bar{C}M_{qu}C + \tfrac{1}{2}(F(A))^2$$

$$\equiv S^{(cl)} + S''$$

$S^{(cl)}$ is just the instanton action:

$$S^{(cl)} = 8\pi^2/g^2$$

and the matrix notation is a shorthand for the second order expansion of the action in each of the fields:

$$\tfrac{1}{2}A^{(qu)}M_A A^{(qu)} = \tfrac{1}{2}\int[(D_\mu A_\nu^{(qu)})^2-(D_\mu A_\mu^{(qu)})^2+gA_\mu^{a(qu)}\varepsilon_{abc}F_{\mu\nu}^{b,(cl)}A_\nu^{c(qu)}]d^4x$$

$\bar{\psi} M_\psi \psi = \int \bar{\psi}(x)(\not{D}-m)\psi(x)d^4x$ where $\not{D} = \gamma^\mu D_\mu$

$\frac{1}{2}(F(A))^2$ is the gauge fixing term and the ghost matrix represents the usual ghost term. In these expressions the covariant derivative is taken at $A_\mu^{(cl)}$:

$$D_\mu A_\nu^{(qu)} = \partial_\mu A_\nu^{(qu)} + ig[A_\mu^{(cl)}, A_\nu^{(qu)}]$$

If we now introduce collective coordinates γ_i and a corresponding Jacobian factor $J(\gamma)$ into (3.39), Z_1 is given (at the one-loop level) by(72)(74)(75)

$$<0|0>_1 = \frac{\int \prod_i d\gamma_i J(\gamma_i)[dA^{(qu)}][d\psi^+][d\psi][dc^+][dc]\exp\{-\frac{8\pi^2}{g^2}\}\exp\{-S''\}}{<0|0>_o}$$

$$= N\int \prod_i d\gamma_i J(\gamma_i)(\text{Det}M_A(\gamma))^{-\frac{1}{2}}(\text{Det}M_\psi(\gamma))(\text{Det}M_{gh}(\gamma))\exp\{-\frac{8\pi^2}{g^2}\}$$

(i=1...8 for SU(2))

where $N=(\text{Det}M_A)_o^{\frac{1}{2}}(\text{Det}M_\psi)_o^{-1}(\text{Det}M_{gh})_o^{-1}$ where the subscript o means the determinants are evaluated at $A^{(cl)}=0$. It should be noted that the γ_i are not time dependent so these are normal integrations.

The determinants are as before evaluated by solving the eigenvalue equations: $M\psi_n = E_n\psi_n$ and $\text{Det } M = \prod_n E_n$ (e.g.)

$$M_\psi \psi_n = (\not{D}-m)\psi_n = E_n\psi_n$$

and

$$\text{Det } M_\psi = \prod_n E_n$$

In the gauge field products zero eigenvalues are excluded, since they are included in the collective coordinate integrations.

For SU(2) the collective coordinate integrations over the three global gauge orientations yield a simple numerical factor and one is left with an integration over instanton positions and scale sizes. Writing the instanton in 'tHooft's notation

$$A_\mu^{a(cl)} = \frac{2\eta_{a\mu\nu}(x-z)^\nu}{g((x-z)^2 + \rho^2)}$$

with the instanton position parameter Z explicitly displayed, the γ_i integrations are

$$\int \prod_i d\gamma_i \; J(\gamma_i) = C_o \int d^4z \, d\rho \; J(\rho)$$

where C_o is the numerical factor associated with the global gauge integrations and is basically the volume of the gauge group. Now, since the fermion determinant is independent of $A^{(qu)}$ at the one loop level, the A and ghost determinant may be evaluated independently of the fermion determinant. These determinants contain infinities which must be renormalized. The net result at this one loop level is to replace the coupling g by an effective coupling $\bar{g}(\rho)$ which satisfies the renormalization group equation[56][77]

$$\frac{d\bar{g}}{d(\ln \rho)} = -\beta(\bar{g})$$

$$\Rightarrow \text{for small } \rho \qquad \bar{g}^2(\rho) \sim 1/\ln(\rho\mu)$$

$$\mu = \text{mass scale}$$

The final result for Z_1 is:

$$\langle 0|0\rangle_1 = KV \int \frac{d\rho}{\rho^5} \left(\frac{8\pi^2}{\bar{g}^2}\right) \exp\left\{-\frac{8\pi^2}{\bar{g}^2}\right\} \Delta(m,\rho) \qquad (3.40a)$$

where $K = \frac{1}{4\pi^2} e^{-\alpha(1)}$; $V = 8\pi^2/3$ = result of z integrations†

and $\Delta(m,\rho) = \frac{(\mathrm{Det}\ M_\psi)}{(\mathrm{Det}\ M_\psi)_0}$

The SU(2) instanton may be trivially embedded in a higher SU(N) gauge group to yield an SU(N) instanton[69]. The analagous expression to (3.40a) for SU(N) is[74]:

$$<0|0>_1 = \frac{4V}{\pi^2} \frac{\exp\{-\alpha(1)-2(N-2)\alpha(\frac{1}{2})\}}{(N-1)!(N-2)!} \int \frac{d\rho}{\rho^5} \left(\frac{4\pi^2}{\bar{g}^2}\right)^{2N} e^{-8\pi^2/\bar{g}^2} \Delta(m,\rho) \tag{3.40b}$$

$\alpha(1)$ and $\alpha(\frac{1}{2})$ are numerical factors numerically evaluated by 't Hooft; $\alpha(\frac{1}{2}) \simeq 0.15$ and $\alpha(1) \simeq 0.44$.

Since non-Abelian gauge theories are asymptotically free, $\bar{g}(\rho)\to 0$ as $\rho\to 0$, hence the ρ integration in equation (3.40) will converge for small ρ. Infrared effects, however, cause $\bar{g}$ to be poorly defined at large ρ and thus the large ρ portion of the integration in (3.40) remains ambiguous until the infrared behaviour of non-Abelian gauge theories is well understood. Large ρ also corresponds, though, to the region where the semi-classical, stationary phase approximation will no longer be valid[56].

The fermion determinant $\Delta(m,\rho)$ has been computed for small ρ[72] and for large ρ[76]:

---

†Ore, reference (73)- the finite result for the volume integration is due to the boundary condition of the instanton effectively compactifying spacetime to the O(5) hypersphere.

$$\Delta(m\rho) = \begin{cases} 1.338(\rho m) & \rho m << 1 \quad \dagger \\ 1 - 0.16(\rho m)^2 & \rho m >> 1 \end{cases}$$

The vacuum expectation value of some operator $\hat{O}(A,\psi)$ including the presence of a single instanton may be written in the form

$$<0|\hat{O}(A,\psi)|0>_1 = \frac{\int[dA_\mu][d\psi^+][d\psi]\hat{O}(A,\psi)\exp\{-S[A_\mu^{(cl)}+A_\mu^{(qu)},\psi]\}}{<0|0>} \tag{3.41a}$$

where ghost and gauge fixing terms have been omitted: both numerator and denominator include 0+1 instanton contributions. At the one loop level this becomes

$$<0|\hat{O}(A,\psi)|0> = <0|\hat{O}(A,\psi)|0>_o + KV\int dz^4 \frac{d\rho}{\rho^5}\left(\frac{8\pi^2}{\bar{g}^2}\right)\exp\{-\frac{8\pi^2}{\bar{g}^2}\}\Delta(m\rho)\{<\hat{O}>_{A^{cl}} - <\hat{O}>_o\} \tag{3.41b}$$

Here $<\hat{O}>_{A_{cl}}$ is the expectation value of the operator in the field of a single instanton and will in general be ρ and z dependent (and also global group orientation dependent, which we will assume averaged over(76)) and $<0|\hat{O}|0>_o \equiv <\hat{O}>_o$ is the (normalized) vacuum expectation value of the operator excluding instanton contributions.

The effects of multi-instanton (and anti-instanton) contributions may be included if one approximates the contribution of $n_+(n_-)$ instantons (anti-instantons) by the sum of $n_+(n_-)$ one instanton (anti-instanton) contributions.

---

† The value of the determinant here excludes a contribution which was implicitly included in (5.2) in replacing g by $\bar{g}$, e.g. compare with (76).

Since the sum of two single instanton fields does not satisfy the Yang-Mills field equations, such an approximation can be expected to be valid when the action for the sum of the n fields is not much greater than the action for the exact instanton solution $\frac{8\pi^2 n}{g^2}$. The difference may be considered a form of interaction energy between the instantons and hence the approximation becomes accurate for isolated, non-interacting instantons only. For this reason it is called the dilute gas approximation (D.G.A.)[56][76]. It is convenient to introduce a quantity $D(\rho)$ which may be interpreted as a sort of instanton density[56]:

$$D(\rho) = K\left(\frac{8\pi^2}{\bar{g}^2(\rho)}\right)^4 \exp\left\{-\frac{8\pi^2}{\bar{g}^2(\rho)}\right\} \qquad \text{for } SU(2)$$

Equation (3.40) may thus be rewritten

$$\langle 0|0\rangle_1 = V\int \frac{d\rho}{\rho^5} D(\rho)\,\Delta(m\rho)$$

The vacuum expectation value of an operator $\hat{O}(A,\psi)$, including multi-instanton contributions may now be formally written in the D.G.A. as[76]:

$$\langle 0|\hat{O}(A,\psi)|0\rangle = \frac{\sum\limits_{n_+,n_-}^{\infty} \frac{1}{n_+!\,n_-!}\int \prod\limits_{i=1}^{n_+,n_-} \frac{dz_i\,d\rho_i}{(\rho_i^{\pm})^5} D(\rho_i^{\pm})\,\Delta(m\rho_i^{\pm})\,\langle\hat{O}\rangle_{A^{(cl)}_{n_+n_-}}}{\sum\limits_{n_+,n_-=0}^{\infty} \frac{1}{n_+!\,n_-!}\int \prod\limits_{i} \frac{dz_i^{\pm}\,d\rho_i^{\pm}}{(\rho_i^{\pm})^5} D(\rho_i^{\pm})\,\Delta(m\rho_i^{\pm})} \tag{3.42}$$

where

$$A^{(cl)}_{\mu,n_+,n_-} = \sum_{i=1}^{n_+} A^{(cl)i}_{\mu}(z_i,\rho_i) + \sum_{i=1}^{n_-} \bar{A}^{(cl)i}_{\mu}(z_i,\rho_i)$$

is a sum of single instanton and anti-instanton fields.

3.6 Fermion Zero Modes

The form of the fermion determinant $\Delta(m\rho)$ is proportional to $m\rho$ for small ρ. This implies that, for an identically vanishing fermion mass, the determinant would be zero and the fermion operator M_ψ would develop a zero eigenvalue. This would give a zero value for Z_1 at the one loop level. In other words, at the one loop level, instanton effects disappear when massless fermions are present.

The existence of zero eigenvalues for M_ψ when massless fermions are coupled to a self-dual instanton field can be rigorously proven using the Atiyah-Singer index theorem. This theorem is a modern result from pure mathematics which relates the analytical index of a differential manifold to the topological index. The analytical index is an integer associated with elliptic operators on the manifold and the topological index is defined in terms of the Chern numbers of the manifold.

More specifically, for one elliptic operator L and its adjoint $L^\dagger$, the analytical index is the difference between the number of normalizable zero eigenvalues of $L(n_-)$ and $L^\dagger(n_+)$:

$$I = n_- - n_+$$

Now, it is possible to choose a representation of the Dirac γ matrices so that the massless Dirac equation in a Yang-Mills field takes a form suitable for the application of the index theorem[38]. Explicitly, the massless Dirac equation

$$M_\psi \psi = \gamma^\mu (i\partial_\mu - A_\mu)\psi = 0 \qquad (3.43)$$

becomes, for a choice of γ matrices:

$$\gamma^\mu = \begin{bmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{bmatrix} ; \quad \gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

the two decoupled equations

$$\sigma^\mu (i\partial_\mu - A_\mu)\psi^- = 0$$

$$\bar{\sigma}^\mu (i\partial_\mu - A_\mu)\psi^+ = 0$$

with ψ of (6.1) given by:

$$\psi = \begin{pmatrix} \psi^- \\ \psi^+ \end{pmatrix}$$

and ψ^- and ψ^+ states of definite chirality. With the identification

$$L = \sigma^\mu (i\partial_\mu - A_\mu) \qquad (3.44a)$$

$$L^\dagger = \bar{\sigma}^\mu (i\partial_\mu - A_\mu) \qquad (3.44b)$$

the index theorem states:

$$n_- - n_+ = \frac{1}{16\pi^2} \int d^4x \, \mathrm{Tr}({}^*F^{\mu\nu} F_{\mu\nu}) \quad (F^{\mu\nu} \text{ hermitian}) \qquad (3.45)$$

That is, the difference of the number of zero eigenvalues of L and $L^\dagger$ is just the Pontryagin number or Chern number of the manifold. For a self-dual A_μ it can be shown that only one of n_- or n_+ is non-zero[38]. The zero eigenvalue equation (3.43) therefore reduces to either

$$L\psi^- = 0 \quad \text{or} \quad L\psi^+ = 0$$

For a single instanton, q=1 and (3.45) tells us that the fermion operator M_ψ has exactly one zero eigenvalue and hence $\Delta(m\rho)$ is identically zero as expected from 't Hooft's calculation[72].

The existence of zero modes of M_ψ for massless fermions can also be traced to the existence of an anomaly in the axial vector current operator[45][78][79]. The axial vector current is the current associated with the symmetry (for massless particles):

$$\psi \to e^{i\alpha\gamma_5}\psi \qquad \alpha \text{ a parameter}$$

If one considers the complete eigenvalue equation:

$$M_\psi \psi_n = E_n \psi_n$$

$$\text{or} \qquad \gamma^\mu(i\partial_\mu - A_\mu)\psi_n = E_n\psi_n \tag{3.46}$$

then the axial vector current may be expressed in terms of the ψ_n[38][79]

$$J^\mu_{5(x)} = \text{Tr } i\gamma^\mu\gamma^5 \sum_n \frac{\psi_n(x)\psi_n^\dagger(x)}{E_n + i\mu} \qquad \mu = \text{mass term}$$

$$= \text{Tr } i\gamma^\mu\gamma^5\, S_R(x,x)$$

where S_R satisfies

$$[i\gamma^\mu(\partial_\mu + iA_\mu) + i\mu]S_R(x,y) = \delta^4(x-y)$$

This current is the axial current for a theory with massive fermions and S_R is the corresponding propagator. Now in the process of regularization and renormalization, J^μ_5 acquires an anomaly and its divergence is not the naive term $\mu i 2J_5$. Rather it is given by[80]:

$$\partial_\mu J_5^\mu = 2i\mu J_5(x) + \frac{1}{8\pi^2} \mathrm{Tr}\,({}^*F_{\mu\nu}F^{\mu\nu}) \qquad (3.47)$$

where $$J_5(x) = \sum_n \frac{\psi_n^\dagger(x)\gamma_5\psi_n(x)}{E_n+i\mu}$$

ψ_o can be chosen an eigenstate of γ_5 whereas for ψ_n (eigenvalue E), $\gamma_5\psi_n$ has eigenvalue -E. This follows from multiplying equation (3.46) by γ_5 and noting γ_5 anti-commutes with γ_μ. This then implies:

$$\int d^4x\, \psi_n^\dagger(x)\gamma_5\psi_n(x) = 0 \qquad n\neq 0.$$

$$\int d^4x\, \psi_o^{\dagger(\pm)}(x)\gamma_5\psi_o^{(\pm)}(x) = (\pm)1$$ where $\pm$ denotes positive or negative chirality

Equation (3.47) may now be integrated and only the ψ_o terms will survive in $J_5(x)$ to give $2(n_+-n_-)$. Assuming J_5^μ vanishes on the boundary of 4-space gives

$$n_- - n_+ = \frac{1}{16\pi^2}\int d^4x\, \mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu}) = q$$

That is, the anomalous divergence of the axial vector current, implies this version of the Atiyah-Singer index theorem (or vice-versa). Equation (3.47) may also be considered as a local version of the index theorem.

It should also be noted that when considering only the strictly massless case, the limit $\mu\to 0$ should be taken in (3.47) and when integrating J_5^μ can in general no longer be assumed to vanish, i.e. (3.47) becomes

$$\int d^4x(\partial^\mu J_{\mu 5}) = \frac{1}{8\pi^2}\int d^4x \mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu}) = 2q \qquad (3.48)$$

or $$\oint J_{\mu 5}\, d\sigma^{\mu} = 2q$$

where ∂ denotes boundary of the physical system under consideration.

CHAPTER IV

APPLICATIONS

4.1 Introduction

The analysis of the preceding chapters has concentrated on the formal properties of solutions to non-linear field equations with non-trivial topology. There has, however, been no solid connection made with realistic physical theories. To this point therefore, there is no reason to suppose that such field configurations are any more than mathematical curiosities.

In this chapter some possible physical implications of the existence of such configurations will be outlined. In particular, the role of instantons in non-Abelian gauge theories of elementary interactions will be explored. Most of the effects described here are highly speculative and no real consensus exists as to the correct approach to incorporating instanton effects. For this reason, the main arguments will be very briefly outlined for the major effects thought to originate from instantons.

4.2 The Vacuum in Non-Abelian Gauge Theories

As mentioned in section 2.1, the existence of instanton solutions has important implications for the definition of the vacuum in non-Abelian gauge theories. Before the discovery of the instanton, the vacuum of non-Abelian gauge theories was assumed to be analogous to that in Q.E.D. That is, a unique vacuum state could be defined as a superposition of vacua, $|g\rangle$ ($\Psi[A^g]$), gauge equivalent to the trivial state with $A_\mu = 0$, i.e.

$$|0\rangle = \int [dg]|g\rangle$$

or

$$\Psi_{vac} = \int [dg]\Psi[A_\mu^g]$$

where the gauge group integrations are, as in section 3.3, over infinitesimal gauge transformations. Specifying a physical gauge condition or utilizing the Fadeev-Popov procedure of section 3.3, then yielded a well defined, unique vacuum.

With the advent of instanton solutions to pure non-Abelian gauge theories, it became evident that this was an incomplete definition of the vacuum. The true vacuum also consists of pure gauge fields with $g(x)$ not obtainable continuously from $g(x)=1$. That is, the vacuum consists of pure gauge fields separated into topologically distinct sectors, which may be labelled by the integers.

This may be seen most easily in the gauge $A_o=0$[†]. Considering the vacuum at some fixed time, implies

$$A_i(\vec{x}) = ig^{-1}(\vec{x})\partial_i g(\vec{x}) \qquad i=1,2,3 \qquad (4.1)$$

The only $g(\vec{x})$ which are physically relevant are those which satisfy:

$$g(\vec{x}) \xrightarrow[|\vec{x}|\to\infty]{} 1 \qquad (4.2)$$

[†]A discussion in more general gauges is more complex, e.g.(83).

This follows since those $g(\vec{x})$ which do not satisfy this condition give $A_i(\vec{x})$ which are separated by an infinite energy barrier from the normal vacuum field $A_i=0$. That is, such field configurations would have $A_i \neq 0$ over all space and therefore transition amplitudes from the trivial vacuum configurations with $g(\vec{x})=1$ would involve $\dot{A}_i \neq 0$ over an infinite volume. Such infinite energy transitions have zero amplitude and therefore configurations not satisfying (4.2) may be ignored.

To topologically classify the remaining configurations, one may, as before, consider $g(\vec{x})$ as a mapping from space (now 3-space) into the group manifold. The condition (4.2), however, effectively identifies all points at infinity for the purpose of classifying the mapping $g(\vec{x})$. Now, a 2-dimensional surface with the boundary points identified is topologically equivalent to a sphere S^2 and in the same way, 3-space with the boundary points identified is topologically equivalent (homeomorphic) to a sphere S^3. Therefore, the $g(\vec{x})$ are mappings from $S^3 \to S^3$ (for SU(2)). There are therefore n topologically distinct configurations classified by $\Pi_3(SU(2))=Z$ and hence n topologically distinct vacua $|n\rangle$.

In Chapter II (equation 2.9), it was noted that the number of times the group manifold was covered for each covering of space was given by the wrapping number:

$$n=\frac{1}{24\pi^2}\int d^3x\,\epsilon_{ijk}\mathrm{Tr}\{(g^{-1}(\vec{x})\partial_i g(\vec{x}))(g^{-1}(\vec{x})\partial_j g(\vec{x}))(g^{-1}(\vec{x})\partial_k g(\vec{x}))\}$$

$$= \frac{1}{24\pi^2}\int d^3x\,\epsilon_{ijk}\mathrm{Tr}(A_i A_j A_k) \tag{4.3}$$

An example of $g(\vec{x})$ which gives n=1 is[84]

$$g_1(\vec{x}) = \frac{(\vec{x})^2 - \rho^2}{(\vec{x}^2) + \rho^2} - \frac{2i\rho\vec{\sigma}\cdot\vec{x}}{(\vec{x})^2+\rho^2} \qquad (4.4a)$$

which gives

$$\vec{A}_1(\vec{x}) = ig_1^{-1}\vec{\nabla}\, g_1(x) = -\frac{2\rho}{(\vec{x}^2)+\rho^2}[\vec{\sigma}(\rho^2-(\vec{x})^2)+2\vec{x}(\vec{\sigma}\cdot\vec{x})+2\rho\vec{x}x\vec{\sigma}] \qquad (4.4b)$$

A $g(\vec{x})$ which gives n for equation (4.3) is

$$[g_1(\vec{x})]^n \equiv g_n(\vec{x})$$

Vacuum configurations with differing n are separated by a finite energy barrier. This may be demonstrated by multiplying (4.4b) by a factor ($\frac{1}{2}-\alpha$) and allowing α to vary continuously from $-\frac{1}{2}$ to $+\frac{1}{2}$. F_{ij} is now no longer zero and the energy

$$E = \frac{1}{8}\int d^3x \mathrm{Tr}(F_{ij}F^{ij}) \sim (\alpha^2 - \tfrac{1}{4})^2$$

This shows that the energy has a barrier form between n=0 and n=1, and similarly between other n. This is illustrated schematically in Figure 10.

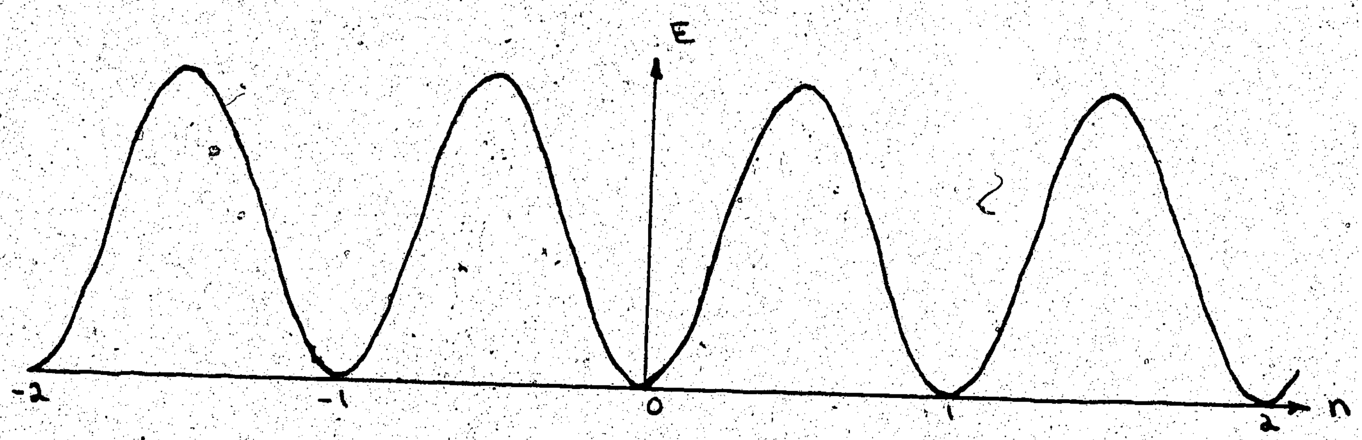


Figure 10. Periodicity of the vacuum in non-Abelian gauge theories.

The understanding of the topological nature of the gauge theory vacuum allows a simple interpretation of the instanton(84)(85). The instanton is a configuration which interpolates between vacua with n differing by 1. This is a semiclassical effect which has no simple description in Minkowski space but in Euclidean space takes the form of a classical phenomenon. By gauge transforming to $A_o=0$, the instanton takes the form

$$A_i(x_4=-\infty) = 0$$

$$A_i(x_4=+\infty) = A_i^1 = ig_1^{-1}\nabla_i g_1$$

Also, in this gauge, it can be seen that:

$$q = \frac{1}{16\pi^2} \int \mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu})d^4x$$

$$= \frac{1}{24\pi^2} \int_{x_4=+\infty} d^3x \mathrm{Tr}(A_i A_j A_k)\varepsilon_{ijk} - \frac{1}{24\pi^2} \int_{x_4=-\infty} d^3x \varepsilon_{ijk} \mathrm{Tr}(A_i A_j A_k)$$

$$= n(x_4=+\infty) - n(x_4=-\infty) = \Delta n \quad (= 1 \text{ for instanton})$$

Consequently, multi-instanton configurations represent tunnelling between vacua with n differing by q.

Since vacua with $n\neq 0$ are connected with the normal vacuum by instantons, the unique vacuum must be considered a superposition of vacua with different n.

$$|0> = \sum_{n=-\infty}^{+\infty} c_n |n> \qquad (4.5)$$

This vacuum should be invariant under "large" gauge transformations, that is, those which change n. This requirement forces c_n to be

$$c_n = e^{in\theta}$$

and under a large gauge transformation (e.g. g_1)

$$|0\rangle_\theta \xrightarrow[g_1]{} e^{-i\theta}|0\rangle_\theta \qquad 0\leq\theta\leq 2\Pi \ .$$

The phase angle θ emerges as an arbitrary parameter characterizing the vacuum, which cannot be determined a priori by any apparent means.

The vacuum to vacuum amplitude may now be written in Euclidean space:

$$\begin{aligned} {}_\theta\langle 0|0\rangle_\theta &= \sum_{n,m} e^{i(n-m)\theta}\langle m|e^{-HT}|n\rangle \\ &= \sum_\nu e^{-i\nu\theta}\int[dA_\mu]\exp\{-\int d^4x L(A(x))\} \text{ for } \nu=m-n \text{ and } T\to\infty \\ &\equiv \sum_\nu\int[dA_\mu]\exp\{-\int d^4x[L(x)+L_\theta(x)]\} \end{aligned}$$

where $\quad L_\theta(x) = \dfrac{i\theta}{16\pi^2}\,\mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu})$

and the path integrals are over A_μ which carry topological charge ν for each term in the sum over ν. The net effect of the new vacuum structure is therefore an effective Lagrangian term proportional to θ. Now, ${}^*F_{\mu\nu}F^{\mu\nu}$ is not P,T invariant so for θ≠0, the theory will in general not conserve parity or time reversal invariance. For Q.C.D., therefore, θ must be very nearly zero. For weak interactions, however, such a term may provide an explanation for observed P and T violations[86].

4.3 The U(1) Problem

If one begins a study of Q.C.D. by ignoring weak interactions, all quark flavors may be assumed to have the same mass. If the mass is initially assumed to be zero, then for N flavors, the theory possesses a chiral SU(N)×SU(N) symmetry. For simplicity, one can assume only two quarks and therefore chiral SU(2)×SU(2) is the relevant symmetry.

The masslessness of the quarks also implies, however, a chiral U(1)×U(1) symmetry. The successes of P.C.A.C.(87) lead one to believe that the SU(2)×SU(2) symmetry is dynamically broken to SU(2) with the triplet of pseudo-scalar pions the Goldstone bosons required by Goldstone's theorem(88). The fact that these pseudo-scalar pions are not massless is accounted for by adding a flavor breaking mass to the quarks in the initial Lagrangian.

Similarly, the U(1)×U(1) symmetry should be broken to U(1), otherwise hadrons would appear in degenerate parity doublets. Standard current algebra techniques show that the corresponding Goldstone boson should have a mass less than $\sqrt{3}\, m_\pi$(89). Unfortunately, no such pseudo-scalar boson exists. This constitutes the U(1) problem circa 1975 (note also (90)).

Since Goldstone's theorem requires a current satisfying $\partial_\mu J^\mu = 0$, a possible solution is that the axial U(1) current J_5^μ is not conserved as naively expected. As was noted in section 3.5, this is in fact the case since the axial current possesses an anomaly and

$$\partial_\mu J_5^\mu = \frac{N}{8\pi^2}\,\mathrm{Tr}({}^*F_{\mu\nu}F^{\mu\nu}) = \frac{N}{8\pi^2}\,(\partial_\mu K^\mu)$$

with K^μ given by equation (2.8). This does not, however, solve the U(1) problem since a new current may be defined

$$\tilde{J}_5^\mu = J_5^\mu - \frac{NK^\mu}{8\pi^2}$$

(N=2 in the case of 2 flavors)

which is conserved:

$$\partial_\mu \tilde{J}_5^\mu = 0$$

Since the new current satisfies the same commutators with quark mass terms as J_5^μ, the U(1) problem reappears as the question of the non-existence of a pseudo-scalar boson associated with the spontaneous breaking of $\tilde{J}_5^\mu$ symmetry.

The generator associated with the current $\tilde{J}_5^\mu$ is

$$Q_5 = \int d^3x\ \tilde{J}_5^o(x)$$

and since Q_5 generates a symmetry

$$[Q_5, H] = 0$$

This symmetry is different than those generated by the other chiral charges, however. A chiral U(1) transformation generated by Q_5, has the following effect on a θ vacuum[84][85]

$$e^{i\alpha Q_{5/2}}|0\rangle_\theta = |0\rangle_{\theta+\alpha} \qquad (4.6)$$

The U(1) chiral charge thus changes θ vacua. Normally an expression such as (4.6) would signal the Goldstone theorem and a corresponding massless mode. The θ symmetry is in this case, however, not a physical symmetry of the theory and a careful analysis of corresponding Ward-Takahashi identities shows that the massless mode has positive and negative metric components in the corresponding Hilbert space and those components cancel for all gauge invariant quantities (e.g. (53)(40)(90)).

U(1) chiral symmetry is thus seen to be broken by the vacuum without a corresponding Goldstone mode and therefore when SU(2)×SU(2) symmetry is broken, the chiral U(1) symmetry does not imply a pseudo-scalar meson in the same manner as chiral SU(2) breaking implies the pion†.

The above result depended on the character of the θ vacua found by interpreting instantons as tunnelling configurations. In Chapter III, however, it was shown that the Atiyah-Singer index theorem implies zero tunnelling amplitude, in the one-loop approximation, when massless fermions are present. It is therefore not apparent that the θ-vacua are the correct vacua for a theory incorporating massless quarks. The index theorem can be avoided however by considering tunnelling configurations (instantons) followed by anti-tunnelling configurations (anti-instantons) so that the net topological charge is zero. Since any tunnelling at all between the n vacua requires a superposition into a θ vacuum, these "virtual" tunnellings are sufficient

†The adequacy of this solution is not yet completely clear, e.g. (91)(92).

to retain our vacuum structure of section 4.2. A more complete treatment of this question can be found in references (53), (84), (85), (91).

4.4 Mass Generation by Instantons

It was mentioned in the last section that the SU(N)×SU(N) symmetry of massless Q.C.D. is expected to be dynamically broken to SU(N) to correspond with the successes of current algebra phenomenology. Recent calculations utilizing the dilute gas approximation indicate that instantons may be responsible for this symmetry breaking by dynamically generating an effective quark mass ((96);(93)-(96)). In these treatments it is assumed that the mass generation effects occur at distances small compared to scales where quark confinement is expected and hence the two problems can be treated independently.

To consider mass generation effects in the dilute gas approximation we begin by considering the propagation of a massive fermion in the presence of a single instanton[(95)(94)] (97).

$$S_F^{(1)} \equiv \langle x| \frac{i}{\not{\partial} + im} |y\rangle_1$$

Comparison with equation (3.41b) of the last chapter shows

$$S_F^{(1)} = S_F^o + K\int \frac{d\rho\, d^4z}{\rho^5}\left(\frac{8\pi^2}{\bar{g}^2}\right)^4 e^{-8\pi^2/\bar{g}^2}\,\Delta(m\rho)\{\langle S_F\rangle_{A^{cl}} - S_F^o\}$$

$$\equiv S_F^o + S_F^{I(+)} \qquad (4.7)$$

where S_F is the normal, free propagator:

$$S_F^o = \langle x \,|\frac{i}{\not\partial+im}|\, y\rangle \equiv \langle 0|T(\psi(x)\overline{\psi}(y))|0\rangle_o$$

and
$$\langle S_F\rangle_{A^{(cl)}} \equiv \langle x|\,\frac{i}{\gamma_\mu(\partial^\mu+igA^{\mu(cl)})+im}\,|y\rangle$$

The contribution of an anti-instanton gives a term identical to (4.7) with $\langle S_F\rangle_{A^{(cl)}}$ replaced with $\langle S_F\rangle_{\overline{A}^{(cl)}}$ ($S_F^{I(+)} \rightarrow S_F^{I(-)}$). The total $S_F^{(1)}$ including one instanton and one anti-instanton is thus

$$S_F^{(1)} = S_F^o + S_F^{I(+)} + S_F^{I(-)}$$
$$\equiv S_F^o + S_F^I$$

Expanding to second order in the dilute gas approximation (i.e. second order in equation 3.42) gives S_F including up to two instantons (anti-instantons). The result is just

$$S_F^{(2)} = S_F^o + S_F^I + S_F^I(S_F^o)^{-1}S_F^I$$

where the factor $(S_F^o)^{-1}$ comes from expanding the denominator in (3.42) and ignoring terms of higher than second order in $D(\rho)$. The inclusion of n instantons in the D.G.A. is now clearly:

$$S_F^{(n)} = S_F^o + S_F^I + S_F^I(S_F^o)^{-1}S_F^I + \ldots + S_F^I(S_F^o)^{-1}\ldots S_F^{I(n)}$$
$$= S_F^o\Big(\sum_{(i)=1}^{n} [S_F^I(S_F^o)^{-1}]^{(i)}\Big)$$

The complete propagator, including all numbers of instantons in the D.G.A. is

$$S_F = \frac{1}{(S_F^o)^{-1}-(S_F^o)^{-1}S_F^I(S_F^o)^{-1}} \equiv \frac{1}{(S_F^o)^{-1}+\Sigma}$$

This shows that the instantons generate a fermion self-energy Σ. Σ possesses terms even in γ matrices and also terms odd in γ matrices. The odd terms can be interpreted as a wavefunction renormalization whereas the even terms can be interpreted as a dynamical mass term[95].

Applying the above analysis to the generation of mass for initially massless quarks, of course, runs into the problem of the fermion zero modes which cause S_F^I to vanish[96]. In reference (56), Callan, Dashen and Gross argue that this does not prevent mass generation however. They consider the virtual tunnellings which remain, as instanton-anti-instanton bound states, and argue that these bound states undergo a phase transition to an approximately free gas of instantons and anti-instantons at some value of effective coupling $\bar{g}$, and in this phase the D.G.A. is valid and mass generation results. In (95), mass generation is approached self-consistently with an initial mass term introduced along with a counterterm. Self-consistency then requires instanton effects to cancel the counter-term. Using the consistency equations they find approximate solutions for dynamical mass generation.

4.5 Speculations on Quark Confinement

Quark confinement remains one of the major problems in modern elementary particle physics. Since quark confinement is expected to be a non-perturbative effect in Q.C.D., it was hoped from the beginning of the exploration of topological non-perturbative soliton states that such an approach would explain quark confinement.

One approach is to consider unbroken Q.C.D. (with $\Pi_1(SU(3))=0$) as equivalent via some dynamical mechanism to a theory with $\Pi_1(G) = Z_3$. This then implies the existence of generalized Nielsen-Olesen vortex tubes which can be interpreted as "strings" joining quark pairs or triplets. The problem of quark confinement thus reduces to explaining the relationship between Q.C.D. and the "effective" theory (16)(98)(99).

An alternate approach was iniated by Polyakov[33][34], and extended by Callan, Dashen and Gross[52][56][107][100]. This assumes Euclidean soliton configurations will dominate the Euclidean path integral to generate quark confinement non-perturbatively. Since this approach relies heavily on the D.G.A. it follows the type of approach outlined above.

As a measure of quark confinement one can consider the vacuum expectation value of the non-integrable phase factor[101] around a Euclidean loop (Figure 11(a))[102]

$$\langle 0|P\exp i \oint_L A_\mu dx^\mu|0\rangle = \int [dA] e^{-S} e^{i\oint_L A_\mu dx^\mu}$$

where P represents path ordering[101] and on the right hand side, S represents the gauge field and quark field action. One may subsequently do the quark field integration, summing over all loops, to give the amplitude for a quark to travel over a closed path as in Figure 11b. This amplitude is just proportional to the amplitude for quark pairs to be created (e.g. by an e^+e^- beam) and then subseque y annihilate. The manner in which the integration over loops is damped for large R via the R dependence of equation 4.8 thus determines the amplitude for observing "free" quarks.

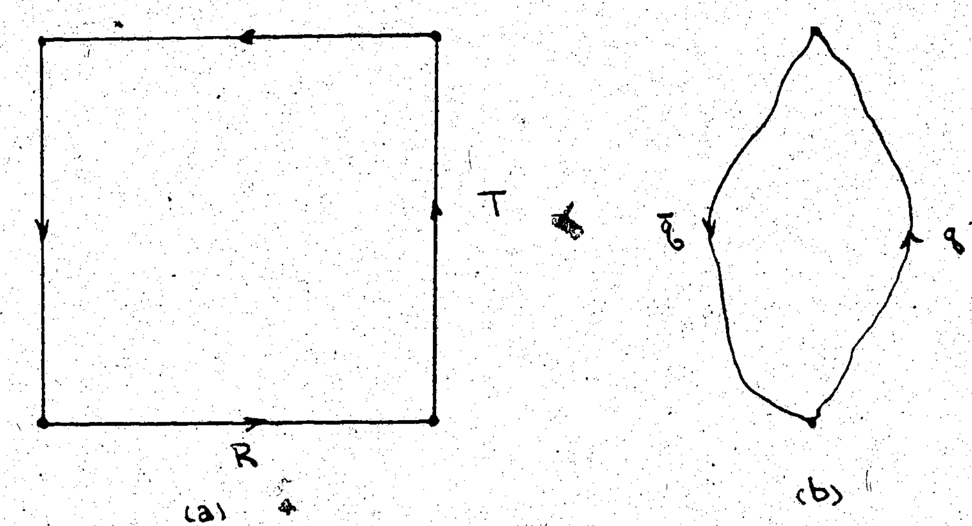


Figure 11. Wilson's Quark Loop.

Working in the gauge $A_0=0$, where only the R integrations contribute to 4.8 (note Figure 11a), one can write

$$<0|PTre^{i\oint A_\mu dx}|0> \xrightarrow[\text{large } T]{} e^{-E(R)T}$$

where E(R) may now be interpreted as the interaction energy of a massive quark-anti-quark pair. If E(R) goes to infinity as $R\to\infty$, then quarks are confined[102].

Instanton effects may be included by calculating equation (4.8) using the dilute gas expansion of equation (3.42) (a perturbation theory calculation gives $E(R) = -\frac{4}{3}(\bar{g}^2/4\pi)R^{-1}$ which is just a Coulomb type interaction energy and is certainly not confining)[60][56]. The net effect of summing the terms in the D.G.A. is to exponeniate the one instanton contribution to equation (4.8)[56]

$$E(R) = -\lim_{T\to\infty} \frac{c}{T} \int dz \frac{d\rho}{\rho^5} D(\rho)\exp\{\oint_L A_I^\mu dx_\mu - 1\}$$

(c=constant)

For large R this approaches a constant and instantons do not appear to confine quarks at this level of approximation[52][56].

Callan, Dashen, Gross have argued, however, that although instantons do not appear to confine quarks, merons (Chapter III) become important in the functional integral and their contribution may possibly create a confining potential. Although the meron configuration written down in Chapter II had infinite action due to the point localization of topological charge, Callan et al. argue that higher order quantum fluctuations will smear this charge over a finite region resulting in finite action for the meron. Configurations with topological charge of one may then be considered as meron bound states with the instanton the minimum energy configuration. The action of a 2 meron configuration is therefore

$$S = \frac{8\pi^2}{g^2} + \Delta S$$

where ΔS can be considered a meron interaction energy and can be estimated to be[56]

$$\Delta S = 6\pi^2/g^2 \ln(R/\sqrt{r_1 r_2})$$

with R the meron separation and $r_1(r_2)$ the radii of the merons.

It is then argued that for a certain value of the effective coupling $\bar{g}$, the merons can be considered as approximately free, and the D.G.A. should be over a meron gas. The application of the D.G.A. is the same as for instantons and the result is[56][100]

$$E(R) \sim R^{7-6\pi^2/\bar{g}^2(R)}$$

For $\bar{g}^2/8\pi^2 \geq 3/28$ this becomes a confining potential and it becomes linear for $\bar{g}^2/8\pi^2 = 1/8$.

This then is the basis for the hope that the quark confinement problem can be solved utilizing Euclidean topological solitons. The physical interpretation given to the confining meron gas[99]-[107], is that a time slice of

Euclidean 4-space shows the meron fields as that of a gas of (approximate) color magnetic monopoles. This gas is then a kind of color magnetic superconductor with hadrons representing bubbles of "normal" vacuum where merons are bound into instantons and gluons can propagate freely. These "bubbles" which represent hadrons are expected to be the field theoretic basis for the phenomenologically successful M.I.T. bag model.

Although this is quite an attractive approach to the solution of the confinement problem, it relies heavily on several approximations which may turn out to be invalid. In particular, the method is inherently semi-classical and ignores the effect of higher order quantum effects. In fact, Witten[(102)] has recently argued, on the basis of the 1/N expansion, that higher order effects should "smear out" the topological charge of an instanton or meron gas, so that the D.G.A. is invalid. Of course such arguments apply also to calculation of mass generation mentioned in the preceding section. Witten's arguments are, however, no more than speculations themselves.

4.6 Conclusion

The question whether topological solitons are physically relevant or mere mathematical curiosities remains unresolved. If the correct theory of weak and electromagnetic interactions is based on a compact gauge group with non-trivial $\Pi_2(G/U(1))$, then magnetic monopoles should exist in the theory as perfectly stable particles. Assuming such a theory were to be well established experimentally, however, proof of the existence of magnetic monopoles would still require detection of such an object. Since the minimum mass of such a monopole would be well into the TeV energy range, the experimental verification of this prediction of topology as applied to particle physics remains unlikely in the near future.

There seems to be no question that pure non-Abelian gauge theories do possess instanton solutions. There also seems to be little question that these solutions can be interpreted as tunnelling events between topologically distinct classical vacua. The key question is of course whether these instantons have any true physical significance. The preceding sections of this chapter indicate how several extremely important effects in elementary particle physics could be due to instantons. All these instanton applications, however, suffer from two very serious problems; the semi-classical nature of all calculations done up to this point in time and the ambiguity inherent in the integration over instantons of varying size.

The semi-classical nature of instanton calculations is a serious flaw and it is precisely the higher quantum effects which Witten has argued will destroy the concept of an instanton gas. His analysis is based on another non-perturbative approach, the 1/N expansion, and it appears (although this point is certainly unclear at this time) that these two non-perturbative techniques are incompatible. A complete solution to the question of whether the instanton calculations outlined above are reliable will undoubtedly require calculations to higher orders of quantum fluctuations about instanton and multi-instanton configurations. Some progress has been made and the formalism has been developed to the point that quantum effects about a single instanton can be calculated to any order(104) or to any order within the D.G.A.(103). The calculations that have been done to the two loop level(103) show no qualitatively new effects but the calculations remain within the D.G.A.

The problem associated with the scale integrations in instanton calculations is closely related to the validity of the D.G.A. If the behaviour of the scale integration is such that large scale size instantons contribute very little to the effects considered above, then the D.G.A. should be a valid approximation. In fact the quark confinement and mass generation calculations depend crucially on instantons larger than the scale under consideration having little effect thus giving an effective cutoff to the scale integrations. Such approximations must be dealt with very cautiously however, since the scale integrations are present to restore the scale invari[illegible]st when instantons are introduced into a massless Q.C.D[illegible]s reason, arbitrarily truncating the scale inte[illegible]ons [illegible]uld possibly introduce effects similar to sponta[illegible] breaking scale invariance (mass generation?).

The complete solution of this problem will require calculation of quantum effects about exact multi-instanton configurations as well as a detailed understanding of the scale dependence of the effective coupling for Q.C.D. Although approximate calculations have been made with respect to the former(105), the latter will require a much deeper understanding of Q.C.D.

Of course, the applications of instantons to particle physics could be vindicated by an experimental prediction directly verified. Such a possible experimental effect was provided by the prediction of a new light boson, the axion, as a "natural" explanation for θ being zero in strong interactions(86). The axion, however, has not been observed and although the mechanism of reference (86) has not yet been ruled out, the experimental evidence makes verification of instanton physics in this manner unlikely (see e.g. (40)).

Even if one adopts the (I feel extreme) point of view that instantons will prove to be merely interesting configurations of Euclidean Yang-Mills theories, with no important physical consequences, they have nonetheless provided a fascinating connection with pure mathematics. Relationships such as that between the Atiyah-Singer index theorem and the axial vector anomaly have sparked a great deal of interest among physicists in pure mathematics and vice versa. This mutual interest has already led to a considerably deeper understanding of the Yang-Mills field equations and work is currently being done by both pure mathematicians and theoretical physicists.

FOOTNOTES

	Page
Relativistic field theories may also be constructed which exhibit soliton solutions which are stable by dynamical rather than topological mechanisms (2), but these will not be considered here.	3
Everything in this chapter is purely classical and any terminology borrowed from quantum field theory is used merely for convenience.	4
For a formulation of superconductivity in terms of spontaneously broken phase invariance. See e.g. (10).	16
An exact numerical solution has been found when an extra constraint on the coupling is imposed - reference (11).	22
Two standard mathematical references are (13); careful approaches for physicists may be found in (6), (14), (15) and (16).	28
See section 4.1 for more details on this point.	50
Belavin et al. dubbed the solution the pseudo-particle but 't Hooft introduced the name instanton because of the localized nature of the solution in time. Both terms are commonly used.	58
Although these configurations were first discussed by de Alfaro, Fubini and Furlan, the name meron was coined by Callan, Dashen and Gross; the word meron comes from the Greek word for fraction.	63
For a careful mathematical treatment of integration over function spaces, see (67).	69
The gauge field notations are now those of Chapter I, i.e., A_μ, $F_{\mu\nu}$ hermitian rather than anti-hermition, e.g. $A_\mu = A^a_\mu T^a$, etc.	77
As it turns out (section 4.1) the vacuum itself needs a careful definition for non-Abelian gauge theories.	91
Ore, reference (73) - the finite result for the volume integration is due to the boundary condition of the instanton effectively compactifying space-time to the O(5) hypersphere.	95

	Page
The value of the determinant here excludes a contribution which was implicitly included in (5.2) in replacing g by $\bar{g}$, e.g. compare with ([illegible]6).	96
A discussion in more general gauges is more complex, e.g. (83).	104
The adequacy of this solution is not yet completely clear, e.g. (91)(92).	111

BIBLIOGRAPHY

(1) A.C. Scott, F. Chu and D. McLaughlin, Proc. IEEE 61, (1973) 1443.

(2) R. Friedberg, T.D. Lee, A. Sirlin, Phys. Rev. D13 (1976) 2739; Nucl. Phys. B115 (1976) 1, 32.

(3) G.H. Derrick, Journ. Math. Phys. 5 (1964) 1252.

(4) R. Rajaraman, Phys. Rep. 21C (1975) 229.

(5) R. Jackiw, Rev. Mod. Phys. 49 (1977) 681.

(6) S. Coleman in New Phenomena in Subnuclear Physics (Proc. 1975 Int. School of Physics 'Ettore Majorana') ed. A. Zichichi (New York, Plenum Press).

(7) W. Marciano and H. Pagels, Phys. Rep. 36C (1978) 137.

(8) Phys. Rep. 23C (1976) 237.

(9) H.B. Nielsen and P. Olesen, Nucl. Phys. B61 (1973) 45.

(10) D. Lurié in Quantum Fluids, ed. N. Wiser and D.J. Amit (Gordon and Breach, New York 1970) p. 543.

(11) H.J. De Vega and F.A. Schaposnik, Phys. Rev. D14 (1976) 1100.

(12) E.S. Abers and B.W. Lee, Phys. Rep. 9C (1973) 1.

(13) N.E. Steenrod, The Topology of Fibre Bundles (Princeton Univ. Press, Princeton, N.J., 1951).

S.T. Hu, Homotopy Theory (Academic Press, New York, 1966).

(14) L.J. Boya, J.F. Cariñena, and J. Mateos, Fortschritte der Physik 26 (1978) 175.

(15) H.C. Tze and Z.F. Ezawa, Phys. Rev. D14 (1976) 1006; J. Math. Phys. 17 (1976) 2228.

(16) G. 't Hooft in Particles and Fields, eds. A.N. Kamal and D. Boal (Plenum Press, New York 1978).

(17) Yu.S. Tyupkin, V.A. Fateev and A.S. Shuarts, J.E.T.P. Lett. 21 (1975) 42.

M.I. Monastryrskii and A.M. Perelomov, ibid, p. 43.

(18) D. Finkelstein and J. Rubinstein, J. Math. Phys. 9 (1968) 1762.

(19) H. Georgi and S.L. Glashow, Phys. Rev. Lett. 32 (1974) 194.

(20) G. 't Hooft, Nucl. Phys. B79 (1974) 276.

(21) A.M. Polyakov, J.E.T.P. Lett. 20 (1974) 194.

(22) F.A. Bais and J.R. Primack, Nucl. Phys. B123 (1977) 253.

(23) D. Wilkinson and A.S. Goldhaber, Phys. Rev. D16 (1977) 1221.

(24) P. Goddard and D.I. Olive, Rep. Prog. Phys. 41 (1978) 1357.

(25) Int. J. Theor. Phys., 17 (1978) 235-308.

(26) H. Georgi and S.L. Glashow, Phys. Rev. Lett. 28 (1972) 1494.

(27) M. Prasad and C. Sommerfield, Phys. Rev. Lett. 35 (1975) 760.

(28) A.S. Schwarz, Nucl. Phys. B112 (1976) 358.

(29) S. Arafune, P.G.O. Freund and C.J. Goebel, J. Math. Phys. 16 (1975) 433.

(30) S. Coleman, Comm. Math. Phys. 55 (1977) 113.

(31) H. Pagels, Phys. Lett. 68B (1977) 466.
R. Weder, Comm. Math. Phys. 57 (1977) 161.

(32) A. Jackiw, I. Muzinich, C. Rebbi, Phys. Rev. D17 (1978) 1576.

(33) A.M. Polyakov, Phys. Lett. 59B (1975) 82.

(34) A.M. Polyakov, Nucl. Phys. B120 (1977) 429.

(35) W. Drechsler and M.E. Mayer, Fiber Bundle Techniques in Gauge Theories (Springer-Verlag, Berlin 1977).

(36) B. Schroer in Facts and Prospects of Gauge Theories, ed. P. Urban, Acta Physica Austriaca, Suppl. XIX, 1978, p. 155.

(37) L.S. Pontryagin, Topological Groups (2nd ed., Gordon and Breach, New York, 1977).
R. Gilmore, Lie Groups, Lie Algebras, and some of Their Applications (J. Wiley and Sons, New York, 1974).

(38) R. Jackiw, C. Nohl and C. Rebbi in Particles and Fields, ed. A.N. Kamal and D. Boal (Plenum Press, New York, 1978), p. 199.

(39) A.A. Belavin, A.M. Polyakov, A.S. Schwartz and Yu.S. Tyupkin, Phys. Lett. 59B (1975) 85.

(40) R.J. Crewther in Facts and Prospects in Gauges Theories, Ibid reference (36).

(41) M. Atiyah and R. Ward, Comm. Math. Phys. 55 (1977) 117.

(42) F. Wilczek in Quark Confinement and Field Theory, ed. D. Stump and D. Weingarten (Wiley, New York, 1977); F. Corrigan and D. Fairlie, Phys. Lett. 67B (1977) 69.

(43) R. Jackiw, C. Nohl and C. Rebbi, Phys. Rev. D15 (1977) 69.

(44) G. 't Hooft, Coral Gables Lectures, 1977, unpublished.

(45) L.S. Brown, R. Carlitz and C. Lee, Phys. Rev. D16 (1977) 417.

(46) A. Schwarz, Phys. Lett. 67B (1977) 172.

(47) R. Jackiw and C. Rebbi, Phys. Lett. 67B (1977) 189.

(48) M. Atiyah, N. Hitchin and J. Singer, Proc. Natl. Acad. Sci. USA 74 (1977) 2662.

(49) A.A. Belavin and V.E. Zakharov, Phys. Lett. 73B (1978) 53.

(50) E. Corrigan in Phys. Rep. 49 (1979) 91 at p. 95.

(51) U. de Alfaro, S. Fubini and G. Furlan, Phys. Lett. 73B (1978) 463.

(52) C. Callan, Jr., R. Dashen and D. Gross, Phys. Lett. 66B (1977) 375.

(53) S. Coleman, Erice Lectures, 1977, to be published.

(54) J. Schwinger, Proc. Nat. Acad. Sci. 37 (1951) 452.

(55) J. Iliopoulos, C. Itzykson and A. Martin, Rev. Mod. Phys. 47 (1975) 165.

(56) C.G. Callan, Jr., R. Dashen and D. Gross, Phys. Rev. D17 (1978) 2717.

(57) R.F. Dashen, B. Hasslacher and A. Neveu, Phys. Rev. D10 (1974) 4114; 4136.

(58) J. Goldstone and R. Jackiw, Phys. Rev. D11 (1975) 1486.
A. Klein and F. Krejs, Phys. Rev. D12 (1975) 3112.
L. Jacobs, Phys. Rev. D13 (1976) 2278.

(59) L.D. Fadeev and V.E. Korepin, Phys. Rep. 42C (1978) 1.

(60) E. Tomboulis, Phys. Rev. D12 (1975) 1678.

N.H. Christ and T.D. Lee, Phys. Rev. D12 (1976) 1606.

M. Creutz, Phys. Rev. D12 (1975) 3126.

E. Tomboulis and G. Woo, Ann. Phys. 98 (1976) 1.

(61) C.G. Callan, Jr. and D. Gross, Nucl. Phys. B93 (1975) 29.

(62) L. Leplae, F. Mancini and H. Umezawa, Phys. Rev. B2, (1970) 3594.

H. Matsumoto, N.J. Papastamatiou and H. Umezawa, Nucl. Phys. B82 (1974) 45; ibid. B97 (1975) 90.

H. Matsumoto, P. Sodano and H. Umezawa, Phys. Rev. D19 (1979) 511.

G. Semenoff, H. Matsumoto and H. Umezawa, Univ. of Alta. preprint (1979).

(63) J.L. Gervais and B. Sakita, Phys. Rev. D11 (1975) 2943.

J.L. Gervais, A. Jevicki and B. Sakita, Phys. Rev. D12 (1975) 1038.

(64) J.L. Gervais and A. Jevicki, Nucl. Phys. Rev. B110 (1976) 93.

(65) K.S. Cheng, J. Math. Phys. 13 (1972) 1723.

(66) R.P. Feynman and A.R. Hibbs, Quantum Mechanics and Path Integrals, (McGraw-Hill, New York, 1965).

(67) I. Gelfand and N. Vilenkin, Generalized Functions, Vol. 4, (Academic Press, 1964).

A.A. Slavnov, Theor. Math. Phys. 22 (1975) 177.

(68) J.D. Bjorken, S.D. Drell, Relativistic Quantum Fields, McGraw Hill, New York, 1965.

(69) C. Bernard, N. Christ, A. Guth and E. Weinberg, Phys. Rev. D15 (1977) 2967.

(70) F.A. Berezin, The Method of 2nd Quantization, Academic Press, New York, 1966.

(71) L. Fadeev, Les Houches Lectures, eds. R. Balian and J. Zinn-Justin (North Holland, Amsterdam, 1976).

(72) 't Hooft, Phys. Rev. D14 (1976) 3432; erratum Phys. Rev. D18 (1978) 2199.

(73) A. Belavin, A. Polyakov, Nucl. Phys. B123 (1977) 429.
F.R. Ore, Jr., Phys. Rev. D16 (1978) 2577.
S. Chadha, A. D'Adda, P. DiVecchia and F. Nicodemi, Phys. Lett. 72B, (1977) 103.

(74) C. Bernard, Phys. Rev. D19 (1979) 3013.

(75) A. Jevicki, Nucl. Phys. B117 (1976) 365.

(76) N. Andrei and D.J. Gross, Phys. Rev. D18 (1978) 468.

(77) H.D. Politzer, Phys. Rep. 14C (1974) 129.

(78) N.K. Nielsen, B. Schroer, Nucl. Phys. B127 (1977) 493.

(79) J. Kiskis, Phys. Rev. D15 (1977) 2329.
L.S. Brown, R.D. Carlitz, D.B. Creames and C. Lee, Phys. Lett. B71 (1977) 103; Phys. Rev. D17 (1978) 1583.

(80) Treiman, Jackiw, Gross, Lectures on Current Algebra and its Applications, Princeton Univ. Press, 1972, at p. 97.

(81) L.D. Fadeev and V.N. Popov, Phys. Lett. 25B (1967) 29.

(82) B.W. Lee, Les Houches Lectures, 1975, ibid. (71).
Marciano and Pagels - Q.C.D., ibid. (7), p. 154.

(83) S. Wadia and T. Yoneya, Phys. Lett. 66B (1977) 341.

(84) R. Jackiw and C. Rebbi, Phys. Rev. Lett. 38 (1977) 1440.

(85) C.G. Callan, R.F. Dashen and D.J. Gross, Phys. Lett. 63B (1976) 334.

(86) R.D. Peccei and H.R. Quinn, Phys. Rev. Lett. 38 (1977) 1440.
S. Weinberg, Phys. Rev. Lett. 40 (1978) 223.
F. Wilczek, ibid. 279.

(87) S.L. Adler and R. Dashen, Current Algebras, Benjamin, New York 1968.
J.J. Sakurai, Currents and Mesons, Univ. of Chicago Press, 1969.

(88) Y. Nambu, Phys. Rev. Lett. 4 (1960) 380.
J. Goldstone, Nuovo Cimento 19 (1961) 154.

(89) S. Weinberg, Phys. Rev. D11 (1975) 3583.

(90) J. Kogut and L. Susskind, Phys. Rev. D11 (1975) 3594.
H. Pagels, Phys. Rev. D13 (1976) 343.

(91) R.J. Crewther, Cern Preprint Th-2546, October 1978.
R.J. Crewther, Phys. Lett. 70B (1977).

(92) E. Witten, Nucl. Phys. B149 (1979) 285.

(93) D.G. Caldi, Phys. Rev. Lett. 39 (1977) 121.

(94) R.D. Carlitz and C. Lee, Phys. Rev. D17 (1973) 3238.

(95) R.D. Carlitz and D.B. Creamer, Anns. of Phys. 118 (1979) 429.

(96) J. Hiertarinta, W. Palmer and S. Pinsky, Phys. Rev. Lett. 40 (1978) 421.

(97) L.S. Brown, R.D. Carlitz and D.B. Creamer and C. Lee, Phys. Lett. B71 (1977) 103; Phys. Rev. D17 (1978) 1583.
R. Jackiw and C. Rebbi, Phys. Rev. D14 (1976) 517.

(98) G. 't Hooft, Nucl. Phys. B138 (1978) 1.

(99) F. Englert and P. Windey, Phys. Rep. 49 (1979) 91 at 173.

(100) D. Gross, Phys. Rep. 49 (1979) 91.

(101) T.T. Wu and C.N. Yang, Phys. Rev. D12 (1975) 3845.

(102) K. Wilson, Phys. Rev. D10 (1974).

(103) H. Levine and L.G. Yaffe, Phys. Rev. D19 (1979) 1225.

(104) D. Amati and A. Rouet, Il. Nuovo Cimento 50A (1979) 265.

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

(106) E. Witten, Nucl. Phys. B149 (1979) 285.

(107) C. Callan, Jr., R. Dashen and D. Gross, Phys. Rev. D19 (1979) 1826.